Molecular Evolution of Equine Influenza Virus Non-Structural Protein 1

Caroline Marie Chauché

Submitted in fulfilment of the requirements for the degree of
Doctor of Philosophy in Virology

Institute of Infection Immunity & Inflammation
College of Medical, Veterinary and Life Sciences
Oilthigh Ghlaschu – University of Glasgow

December 2017

ABSTRACT

Influenza A viruses (IAVs) are common infections of certain avian reservoir species, and they periodically transfer to mammalian hosts. These cross-species jumps are usually associated with sporadic outbreaks, and on rare occasions lead to the establishment of a lineage in the new host species. The immune pressure exerted by the new host on the emergent virus forces it to evolve and adopt strategies to evade immunity in order to survive in nature. Understanding the biological mechanisms that allow successful inter-species transmission and adaptation to mammals is crucial to develop the theoretical tools required to predict and/or control emergence of new viruses in humans and animals. H3N8 equine influenza virus (EIV) represents an interesting model to study the dynamic of within-host variation of an avian-origin IAV. Indeed, this virus has emerged from birds in 1963 and has circulated in horse populations for more than fifty years despite the availability of vaccines. Evidence of evolution of EIV virulence factor non-structural protein 1 (NS1) also exists. NS1 is the main viral antagonist of the host interferon (IFN) response, and it relies on different strategies for overcoming these responses, which varies depending on the viral strain. While some NS1 proteins effectively block the induction of IFN and IFN stimulated genes (ISGs), others block general gene expression at a post-transcriptional level, and therefore reduce the synthesis of IFN and ISGs indirectly. Importantly, little is known about the contribution of these NS1 functions to EIV infection phenotype and adaptation to horses.

In this work, we characterised NS1 proteins spanning the entire EIV lineage and showed that NS1s from different time periods after EIV emergence counteract the IFN response using different and mutually exclusive mechanisms. While EIVs circulating in the early 1960s blocked general gene expression by a NS1-mediated blockade of the cleavage and polyadenylation specificity factor 30 (CPSF30), NS1s from contemporary EIVs specifically inhibit the induction of ISGs by interfering with the JAK/STAT pathway. These contrasting anti-IFN strategies are associated with two mutations that appeared sequentially during EIV evolution, E186K substitution and C-terminal truncation. These changes in NS1 allowed contemporary EIVs to replicate in the presence of high levels of IFN. The results shown here with EIV indicate that the interplay between virus evolution and immune evasion plays a key role in IAV mammalian adaptation.

TABLE OF CONTENTS

TABLES

FIGURES

PUBLICATIONS BASED ON PROJECT RESEARCH

- **Mammalian adaptation of an avian influenza A virus involves stepwise changes in NS1.** **Chauché C.***, Nogales A.*, Ahmad Shanizza A.I., Zhu H., Goldfarb D., Gu Q., Parrish C.R., Martínez-Sobrido L., Marshall J.F., and Murcia P.R. Journal of Virology (2018), 92(5). pii: e01875-17. doi: 10.1128/JVI.01875-17.

CONTRIBUTED TO, BUT NOT INCLUDED IN THIS THESIS

- **The K186E Amino Acid Substitution in the Canine Influenza Virus H3N8 NS1 Protein Restores Its Ability To Inhibit Host Gene Expression.** **Chauché C***, Nogales A*, DeDiego ML, Topham DJ, Parrish CR, Murcia PR, Martínez-Sobrido L. Journal of Virology (2017), 91(22). pii: e00877-17. DOI: 10.1128/JVI.00877-17.

- **Science-in-brief: Clinical highlights from BEVA Congress 2016.** **C. Chauché** and R. Kennedy. Equine Veterinary Journal (2017) 49:10-12. DOI: 10.1111/evj.12644.

- **Temperature-Sensitive Live-Attenuated Canine Influenza Virus H3N8 Vaccine.** Aitor Nogales, Laura Rodriguez, **Caroline Chauché**, Kai Huang, Emma C. Reilly, David J. Topham, Pablo R. Murcia, Colin R. Parrish, Luis Martínez-Sobrido. Journal of Virology (2017) 91:e02211-16. DOI: 10.1128/JVI.02211-16.

- **Canine influenza viruses with modified NS1 proteins for the development of live-attenuated vaccines.** Aitor Nogales, Kai Huang, **Caroline Chauché**, Marta L. DeDiego, Pablo R. Murcia, Colin R. Parrish, Luis Martínez-Sobrido. Virology (2017) 500:1–10. DOI: 10.1016/j.virol.2016.10.008.

ACKNOWLEDGEMENTS

Part of the work presented in this thesis has been carried out by collaborators or members of the Murcia laboratory:

- The NS1/CPSF30 co-immunoprecipitation experiments (Chapter 4) were carried out by Aitor Nogales, Department of Microbiology and Immunology, University of Rochester, Rochester, New York, USA.
- The phylogenetic analysis of EIV NS1s (Chapter 4) was carried out by Henan Zhu, MRC-University of Glasgow Centre for Virus Research.
- The cloning of some NS1 SAM-pCAGGS constructs (Chapter 4) was performed by Azzy Iyzati binti Ahmad Shanizza, master student under my supervision at the MRC-University of Glasgow Centre for Virus Research.
- The bioinformatic analysis of RNA sequencing data (Chapter 6) was executed by Quan Gu, MRC-University of Glasgow Centre for Virus Research.

This Thesis was supported by the Horserace Betting Levy Board (VET/RS/2512) and the University of Glasgow Veterinary Fund Scheme.

As for my personal acknowledgments…

I would like to thank Pablo Murcia for giving me the opportunity to join his laboratory. These three years of PhD in the Murcia lab have been intense in many ways, and I have been forced to grow both professionally and personally. Overall, I consider this scholarship a success.

Most importantly, I would like to say a huge thank you to John Marshall for supporting me throughout this PhD. Thank you for being encouraging, for believing in me even when I was at the lowest. Thank you for never giving up on me, even in the most difficult moment. I will always be grateful, and I wish to become one day as great a mentor as he was to me.

I also feel lucky to have had the opportunity to work in collaboration with Luis Martinez-Sobrido and Aitor Nogales, our collaborators from Rochester University. They were incredibly efficient and at the same time very friendly and generous. Thank you for including me in your research projects and giving me the opportunity to publish with you!

Thank you as well to Mariana Varela, Maxime Ratinier, Filipe Nunes, Ben Hale, Sam Wilson, Edward Hutchinson, Georgia Perona-Wright and Christ Boutell for all the useful discussions, advice and encouragement.

I would to say a big thank you to my super GG, Joanna, Fiona, Sema, Alice, Azzy, Henan, Yasmin and Julien for being such great labmates! Thank you also to other members of my CVR family, Idoia, Virginia, Eleonora, Navapon, Gail, Alice (Broos), Jin, Junje. I had a great time with you all!

Thank you to my family for encouraging me to pursue a scientific career.

And wait, wait, wait! Whenever you're ready… A whopping thank you mi Guanche for being such a great partner (and also for making such a great coffee!). Thank you for supporting me, encouraging me, also telling me off when needed, and most importantly, for making me crack up everyday! You are the best!

DECLARATION

I, Caroline Marie Chauché, hereby certify that, except where explicit reference is made to the contribution of others, this thesis is the result of my own work and has not been submitted for any other degree at the University of Glasgow or any other institution. Some of the work has been published in appropriate journals, as indicated in page 11. I acknowledge that the work undertaken for the completion of this PhD thesis was supported by a Veterinary Research Training Scholarship (VET/RS/2512) awarded to Pablo R. Murcia and John F. Marshall. In addition, this research was partially funded by the Horserace Betting Levy Board (Grant 779) and Veterinary Fund Small Grant Scheme of the University of Glasgow, also awarded to Pablo R. Murcia and John F. Marshall.

Date: 12th of December 2017

Signature: 

Printed name: Caroline Marie Chauché

ABBREVIATIONS

BGH	-	Bovine Growth Hormone
bp	-	Base pair
C-terminal	-	Carboxy-terminal
CDS	-	Coding Sequence
CIV	-	Canine influenza virus
CMV	-	Cytomegalovirus
CPSF	-	Cleavage and Polyadenylation Specificity Factor
cRNA	-	Complementary RNA
cRNP	-	Complementary Ribonucleoprotein
DEG	-	Differentially Expressed Gene
DMEM	-	Dulbecco's Modified Eagle's Medium
DMSO	-	Dimethyl Sulfoxide
DNA	-	Deoxyribonucleic acid
E.Derm	-	Equine Dermal fibroblast
ED	-	Effector Domain
EIV	–	Equine Influenza Virus
ESP	-	Expert Surveillance Panel
F/79	-	A/equine/Fontainbleau1/1979
FC1	-	Florida Clade 1
FC2	-	Florida Clade 2
FF-Luc	-	*Firefly* luciferase
FFU	-	Focus Forming Unit
HAU	-	Hemagglutination Assay Unit
HEK293T	-	Human Embryonic Kidney cells 293
HI	-	Hemagglutination inhibition assay
HSV-TK	-	Herpes Simplex Virus-Thymidine Kinase
IAE	-	Influenza Associated Encephalitis/Encephalopathy
IAV	-	Influenza A virus
IFN	-	Interferon
ISG	-	IFN-Stimulated Gene
ISRE	-	Interferon-Stimulated Response Element
K/02	-	A/equine/Kentucky/5/2002
K/91	-	A/equine/Kentucky/1/1991
K/95	-	A/equine/Kentucky/1/1995
K/99	-	A/equine/Kentucky/1/1999

kb	-	Kilo base
kDa	-	Kilo Dalton
LAIV	-	Live-attenuated Influenza vaccines
LB	-	Luria-Bertani
M/63	-	A/equine/Miami/1/1963
MDCK	-	Madin-Darby Canine Kidney cells
mRNA	-	Messenger RNA
N/03	-	A/equine/Newmarket/5/2003
NCBI	-	National Center for Biotechnology Information
NEP	-	Nuclear Export Protein
NP	-	Nucleoprotein
NS1	-	Non-Structural Protein 1
nSTAT	-	Nuclear STAT
O/03	-	A/equine/Ohio/1/2003
ORF	-	Open-Reading Frame
P-STAT	-	Phospho-STAT
PABPII	-	Poly(A)-Binding Protein II
PAF	-	Polymerase Associated Factors
PAMP	-	Pathogen-Associated Molecular Pattern
PBM	-	PDZ Binding Motif
PBS	-	Phosphate Buffered Saline
pCAGGS	-	CMV enhancer chicken b-Actin promoter rabbit beta-Globin splice acceptor site
PCR	-	Polymerase Chain Reaction
PDZ	-	PSD-95, Discs-large, ZO-1
PL	-	PDZ domain Ligand
Pol	-	Polymerase
PRR	-	Pathogen-Recognition Receptor
PS	-	Penicillin/Streptomycin
RBD	-	RNA-Binding Domain
REN-Luc	-	*Renilla* Luciferase
RIG-I	-	Retinoic Acid-Inducible Gene-I
RNA	-	Ribonucleic Acid
RNA Pol II	-	RNA Polymerase II
S/89	-	A/equine/Sussex/1/1989
SAM	-	Splice Acceptor Mutant
SeV	-	Sendai Virus
SNP	-	Single Nucleotide Polymorphism

SP/69	-	A/equine/Sao Paulo/1/1969
STAT	-	Signal Transducer and Activator of Transcription
TPCK	-	Tosylsulfonyl Phenylalanyl Chloromethyl Ketone
TPM	-	Transcript Per Kilobase Million
U/63	-	A/equine/Uruguay/1/1963
uIFN	-	Universal Interferon
UK	-	United Kingdom
USA	-	United State of America
UTR	-	Untranslated Region
vRNA	-	Viral RNA
vRNP	-	Viral Ribonucleoprotein
VSV	-	Vesicular Stomatitis Virus

Chapter 1

Introduction

1. General introduction

Influenza viruses cause an acute febrile respiratory disease in various host species, including horses. They are responsible for seasonal infections or explosive outbreaks with high morbidity or mortality in susceptible populations.

1.1. Equine Influenza

Equine influenza is a highly contagious disease in horses and other equids, which most commonly affects two- and three-year-old individuals and may be the most common cause of respiratory illness in racehorses (Arthur and Suann, 2011). Some tracks may even experience several influenza outbreaks within a racing season, which lead to significant economic losses for the equine industry (Myers C, 2006, Crawford et al., 2005, Smyth et al., 2011, Park et al., 2004). Indeed, even if EIV infection is generally self-limiting and the majority of horses recover uneventfully, it strongly affects a horse's performance, and requires time out of work for recovery (several weeks to months). Infected horses may also suffer life-threatening complications, such as bacterial pneumonia, particularly when not provided with an adequate period of rest.

1.1.1. Aetiology

The etiologic agent of equine influenza is an Influenza A virus (IAV), which belongs to the family *Orthomyxoviridae*. This family comprises six genera: *Influenzavirus A, B, C, and D, Thogotovirus* and *Isavirus*. The classification system places influenza viruses into genera and species (A, B, C or D), then into 18 hemagglutinin (HA) and 11 neuraminidase (NA) subtypes (*Center for Disease Control and Prevention, Types of Influenza Viruses [2017 December 12] cited http://www.cdc.gov/flu/about/viruses/types.htm*). Two subtypes have been identified in horses: H7N7 and H3N8 (see section 1.1.5). Viruses are further categorized by host, geographical origin, strain, number and year of isolation. Of note, Influenza A viruses can infect horses as well as other host species, such as humans, donkeys, dogs, pigs, birds, mink, seals, camels, cats, and whales (Parrish and Kawaoka, 2005). In contrast, Influenza B viruses are only pathogens of humans, Influenza C viruses can infect pigs and humans, although they rarely cause disease, and Influenza D viruses were recently discovered and are thought to primarily infects cattle (*Center for Disease Control and*

Prevention, Types of Influenza Viruses [2017 December 12] cited http://www.cdc.gov/flu/about/viruses/types.htm).

1.1.2. Clinical signs

EIVs are highly contagious pathogens spread by infectious secretions and aerosols generated by frequent sneezing and coughing. Close contact between individuals seems to be necessary for rapid transmission of the disease, however, fomites may also contribute to virus distribution (Myers C, 2006). In horses, the course of infection usually lasts between two and ten days in uncomplicated cases, and signs of disease appear three to five days following exposure to the virus. Virions are excreted during the incubation period, and individuals remain infectious for at least five days after the onset of clinical disease. In addition, an immunologically naïve horse will shed larger amounts of virus for a longer period of time than a horse previously exposed to viral antigens through vaccination or natural infection (Myers C, 2006).

Infected horses exhibit a sudden fever, which may be biphasic ranging from 39.1°C to 41.7°C, accompanied by serous nasal discharge, anorexia, depression, and in most cases a repetitive harsh, dry, explosive and non-productive cough. Submandibular lymphadenopathy, and pharyngitis, tracheitis, as well as mild bronchitis can be found during endoscopic examination of the respiratory tract, associated with minimal changes in lung sounds (Myers C, 2006, Elton and Bryant, 2011). Some horses exhibit myalgia, myositis, and limb oedema, and are thus reluctant to move. In some cases, a mild form of azoturia (myoglobinuria) may occur. In young foals, influenza is severe and produces signs of viral pneumonia, which may lead to death within 48 hours (Elton and Bryant, 2011, Daly et al., 2011). Similarly to humans, cases of EIV associated with neurological signs have been described and are associated with high levels of pro-inflammatory cytokines (i.e. IL-6 and TNF-α) (Daly et al., 2011).

1.1.3. Diagnosis

As EIV spreads rapidly among susceptible horse populations, an early diagnosis is crucial to prevent an outbreak from occurring. Historically, paired serology samples collected during the acute and convalescent phases of the disease has been the diagnostic technique of choice. Antibody titres can be measured either by hemagglutination inhibition (HI), virus neutralization test, or single radial hemolysis. Serological assays have also been

developed to detect antibodies against the non-structural protein 1 (NS1), which are not produced following classic vaccination regimes (see section 1.3.1 for more information about NS1). This test is particularly useful for differentiating vaccinated from non-vaccinated horses (Myers C, 2006).
Several diagnostic tests are also available, such as an ELISA assay (Directigen™ Flu ELISA, Becton Dickison, Microbiology Systems, Mississauga, Ontario, Canada), or a Polymerase Chain Reaction (PCR), both based on the detection of the viral nucleoprotein (NP) gene (Bryant et al., 2010). Furthermore, immuno-PCR techniques combining ELISA and PCR to detect HA and NS1 are under investigations at present.
Finally, virus isolation can be performed when grown from a nasopharyngeal swab in embryonated chicken eggs or passaged through Madin-Darby canine kidney (MDCK) cells, and detected by HI tests using specific antisera to H3N8 and H7N7 viruses. Results usually take several days to weeks; however, it is the only diagnostic procedure that yields a virus isolate that can be characterized antigenically and genetically. EIV characterization is vital for surveillance, and provides virus for the manufacture or update of vaccines (Myers C, 2006, Damiani et al., 2008)

1.1.4. Prevention and treatment

The increased international movement of horses for breeding and competition purposes has complicated the task of veterinarians and is recognized as an important factor in the spread of EIV throughout the world. A better understanding of the host immune response to natural infection and vaccines, as well as the strategies employed by the virus to escape this response is crucial to elaborate better control measures and more efficient vaccines in the future (Myers C, 2006).

To date, disease prevention remains the best way to protect horses, as the treatment against EIV is mainly symptomatic and vaccines have often failed to provide adequate protection. Individuals demonstrating clinical signs of disease or that have been in contact with affected horses should be placed in quarantine and diagnostic techniques should be conducted. Furthermore, contaminated stables, equipment and vehicles should be disinfected.

The treatment of equine influenza consists in supportive therapies with strict stall rest, good nursing care, access to clean water, electrolyte solutions and palatable food. Antitussives are contraindicated as the cough reflex enhances clearance of the virus, and

bronchodilators are of limited value as the pathophysiology of the disease does not include significant bronchoconstriction in uncomplicated cases. (Myers C, 2006) Antiviral agents, such as Amantadine and Rimantadine, are expensive and of little benefit to an already infected horse but could be valuable to protect horses during an outbreak. Amantadine has been shown to prevent disease in 90% of experimentally infected horses, however, it has also been reported to induce seizures that can be fatal. Rimantadine is a safer and more effective derivative of amantadine (Myers C, 2006). Nonsteroidal anti-inflammatory drugs (NSAIDs) may be administered to counteract fever and subsequent inappetence, but they must be used with caution in dehydrated horses. In addition, antimicrobials (Penicillin G or Trimethoprim/Sulfonamide) may be indicated if the horse is at risk of secondary bacterial infection. Furthermore, to allow a complete recovery and avoid complications, a horse should have one week of rest for each day of fever, thus approximately thirty to sixty days. Returning to work earlier may favour secondary bacterial infections and long-term complications, such as myocarditis and recurrent airway obstruction (Myers C, 2006).

An effective vaccination regime is also critical to protect equine populations, and vaccination became mandatory for racing Thoroughbreds in 1981. EIV vaccines currently licensed for use in the horse are either inert vaccines (i.e. sub-unit vaccines, immune-stimulating complex, DNA vaccines), or live vaccines (live-attenuated virus vaccine and canarypox-based vaccine). A current list of commercialised EIV vaccines can be found at: (Animal Health Trust, equiflunet) http://www.equiflunet.org.uk/.

The United Kingdom (UK) and other countries conduct ongoing surveillance of equine influenza virus, and an Expert Surveillance Panel (ESP) reviews annually all information provided and decide when vaccine strains need to be updated (Daly et al., 2011, Elton and Bryant, 2011). Within the UK, this work is funded by the Horserace Betting Levy Board and carried out by the Animal Health Trust, a World Organisation for Animal Health (OIE) reference laboratory for EIV. Unfortunately, vaccines are rarely updated in accordance with the ESP recommendations, and many current vaccines still contain outdated strains (Gonzalez et al., 2014, Yamazaki and Ichinohe, 2014, Olguin Perglione et al., 2016). EIV vaccines should contain both FC1 and FC2 viruses (see section 1.1.5 and 1.1.6 for more details), such as A/equine/South Africa/04/2003-like or A/equine/Ohio/2003-like viruses for the former, and A/equine/Richmond/1/2007-like viruses for the latter (World organisation for animal health, Equine influenza, [cited 2017 December 12] http://www.oie.int/en/our-scientific-expertise/specific-information-and-recommendations/equine-influenza/).

1.1.5. Epidemiology

Throughout history several introductions of avian IAVs in horses have occurred (Guo et al., 1992), and work done by Kawaoka and colleagues revealed that several genomic segments (PB2, M and NS) were introduced more recently than others (PB1, PA and NP) into equines (Kawaoka et al., 1998) (cf. section 1.1.7.1 for more details about the composition of the Influenza genome). EIVs are evolutionary distinct from other IAVs, and are thought to have been maintained in the equine population for a long time (Webster et al., 1992). Accordingly, Absyrtus, a Greek veterinarian, recorded a disease outbreak in 433 AD resembling influenza in horses, and more recently "The Great Epizootic of 1872" (World organisation for animal health, Equine influenza, [cited 2017 December 12] http://www.oie.int/fileadmin/Home/eng/Media_Center/docs/pdf/Disease_cards/EQUINES-EN.pdf), which spread throughout North America and affected many horses, was likely an outbreak of equine influenza virus, although the subtype has never been identified (Parrish et al., 2015).

EIV was first isolated in 1956 during an outbreak of disease in Prague (Czech Republic). This virus was an H7N7 subtype of avian origin. It was first referred to as equine-1 influenza, and was later identified as A/equine/Prague/1956 (Sovinova and Ludvik, 1959). This virus has been associated with sporadic cases of disease in horses for approximately twenty years post-isolation (Lewis et al., 2011, Hughes et al., 2012).
A different subtype of EIV (H3N8), initially designated as equine-2 influenza, was reported from horses in Florida during an outbreak of disease that started in January 1963 (Scholtens et al., 1964, Waddell et al., 1963). This virus, whose prototype was A/equine/Miami/1/1963, was also of avian origin, and is thought to have originated from South America when Thoroughbred horses exhibiting signs of respiratory disease arrived in Miami (FL, USA) by air from Argentina. Although there was no report of H3N8 EIV circulation in Argentina at that time, there was evidence of circulation and disease in horses in Uruguay, Chile and Brazil (Olguin Perglione et al., 2016, Berrios, 2005). There is no concrete evidence of the origin of this H3N8 EIV, however, Murcia et al. provided evidence that the most recent common ancestor of A/equine/Miami/1/1963 likely arose between 1943 and 1962, and A/equine/Uruguay/1/1963 is thought to be at the origin of this epidemic (Murcia et al., 2011b). H3N8 EIV has spread rapidly among equine populations and caused a major transcontinental pandemic that reached Europe in early 1965 (Paccaud and Paccaud, 1967). This pandemic was distinguished from previous outbreaks (presumably due to H7N7 viruses) by its occurrence in horses of all ages rather than being restricted to younger horses, as was the case until then.

There was an international co-circulation of both H3N8 and H7N7 EIV subtypes in horses between the 1960s and 1970s, and since the 1980s H7N7 EIV has not been isolated in horses, although serological evidence suggests that the virus was still circulating among horses in the early 1990s in Europe and USA (Park et al., 2004, Karamendin et al., 2016). There is no clear explanation of why this subtype disappeared, but it is thought to have been outcompeted by the H3N8 EIV subtype, potentially due to a lower infectivity or transmission efficiency in horses.

For the rest of this work H3N8 EIV will simply be referred to as EIV.

Since viral emergence, H3N8 EIV has been responsible for several outbreaks of disease in vaccinated and naïve horse populations, notably in the UK in 1979, 1989 and 2003, South Africa in 2003, Japan and Australia in 2007, India in 2008-2009, and more recently South America in 2012 (Alves Beuttemmuller et al., 2016, Burrows et al., 1981, Cowled et al., 2009, Ito et al., 2008, Livesay et al., 1993, Newton et al., 2006, Perglione et al., 2016, Virmani et al., 2010, Yamanaka et al., 2008). Today, H3N8 EIV is present in most of the world with the exceptions of New Zealand, Iceland and Australia, which eliminated the virus after the important outbreak of 2007 (World organisation for animal health, Equine influenza, [cited 2017 December 12] http://www.oie.int/fileadmin/Home/eng/Media_Center/docs/pdf/Disease_cards/EQUINES-EN.pdf).
Of note, the latest OIE report of field data on worldwide EIV activity indicated that only individual animal cases and isolated outbreaks of equine influenza had been reported by Ireland, Sweden, the UK and the United State of America (USA). In Europe, there were less equine influenza virus infections than in recent years. Clinically affected horses in the UK were unvaccinated, a single confirmed case in Sweden was of unknown vaccination history, and in Ireland equine influenza cases were confirmed in both vaccinated and unvaccinated horses, but only 10% of horses had up to date vaccination records. In the USA, outbreaks were detected throughout the year with over 30 confirmed cases from 16 states, but no vaccination data were available. Finally, in Asia and South America no equine influenza outbreaks were reported (World organisation for animal health, Equine influenza, [cited 2017 December 12] http://www.oie.int/fileadmin/Home/eng/Media_Center/docs/pdf/Disease_cards/EQUINES-EN.pdf).

As for most IAVs isolated from mammals, the main reservoir of EIV is thought to be wild aquatic birds (Parrish et al., 2015). Avian species have been shown to host multiple IAV strains, most of which are not transmissible to mammals, however, acquisition of point

mutations in the viral genome or segment reassortment among strains can allow IAV to successfully infect and transmit in a new mammalian host population. Cross-species jumps are usually associated with sporadic outbreaks, but on rare occasions emergent viruses manage to establish a lineage in new host species. The compatibility of the viral polymerase with the cellular machinery and the nature of the glycoprotein of surface, HA and NA, are known to be important for successful inter-species jumps (Parrish and Kawaoka, 2005, Parrish et al., 2015, Long et al., 2016), however success or failure of avian-to-mammal transmissions is likely a polygenic trait and the contribution of other viral proteins are only partially understood.

1.1.6. Viral Evolution

The traditional approach to study EIV evolution has been based on HA phylogeny (Bryant et al., 2009, Daly et al., 1996, Damiani et al., 2008, Kawaoka et al., 1989, Lai et al., 2001, Lai et al., 2004), and investigation of vaccine breakdown. The latter has been associated with both inadequately potent H3N8 vaccines and antigenic drift of H3N8 viruses (Bryant et al., 2011, Daly et al., 1996, Mumford, 1999, Newton et al., 2000, Lai et al., 2001, Lai et al., 2004). Antigenic drift refers to the gradual accumulation of genetic mutations, mainly in the highly variable globular head of HA, which causes IAVs to escape recognition by virus neutralizing antibodies and allows it to cause seasonal outbreaks. Another mechanism employed by IAVs to generate genetic diversity and escape the host immune system is the antigenic shift. It occurs when two different IAV strains simultaneously infect the same host cell, and undergo a process called genetic reassortment, where two or more genomic segments are exchanged. During the period of international co-circulation of H7N7 and H3N8 EIV subtypes (1960s and 1970s), heterosubtypic reassortments between H3N8 and H7N7 have occurred, and resulted mainly in H7N7 viruses carrying H3N8 internal genes, notably NP, PB2 and PA segments (see section 1.1.7.1 for more information about the EIV genome). No H3N8 viruses carrying H7N7 internal genes, or H3N7 or H7N8 reassortants have been identified (Bryant et al., 2009).

EIV was initially thought to be genetically stable and the horse was considered a dead-end host for IAVs (Landolt, 2014), but it was later demonstrated that individual EIV infections resulted in an ongoing evolutionary process associated with accumulation of mutations in the viral segments (Murcia et al., 2010). The HA, NA and NS segments were notably identified as sites of localized adaptive evolution (Murcia et al., 2011a). Consistent with this, EIV has circulated as a single phylogenetic lineage during the first twenty years

post-emergence. Then, in the late 1980s it diverged into two antigenically and genetically distinct phylogenetic lineages: the Eurasian and American lineages (Daly et al., 1996, Murcia et al., 2011a), named according to the geographic origin of the isolates. These two lineages co-circulated for a number of years.
In the early 1990s, the American lineage divided into three sub-lineages: 'Kentucky', 'South America' and 'Florida' (Lai et al., 2001), and no virus from the European lineage has been isolated after 1994 (Murcia et al., 2011a).
Finally, in the late 1990's the Florida sub-lineage diverged again into two clades: 'Florida clade 1 (FC1) and 2 (FC2)' (Bryant et al., 2011, Lewis et al., 2011), and ever since FC1 and FC2 viruses have been responsible for influenza cases in horses worldwide. FC1 viruses, whose prototypes are A/equine/Ohio/2003 and A/equine/South Africa/04/2003, are notably endemic in North America, and have also caused major outbreaks in Australia, Japan and South Africa (Cowled et al., 2009, Yamanaka et al., 2008). Small outbreaks have also been reported in Europe between 2007 and 2009 (Bryant et al., 2011, Gildea et al., 2012, Legrand et al., 2013). In contrast, FC2 viruses (prototype: A/equine/Richmond/01/2007) predominate in Asia and Europe, with significant outbreaks reported in China, India and Mongolia, and small-scale outbreaks in multiple countries of Europe (Damiani et al., 2008, Gagnon et al., 2007, Gildea et al., 2013, Landolt, 2014, Virmani et al., 2010, Woodward et al., 2014, Yondon et al., 2013).

Importantly, in the early 2000s EIV has jumped species barriers and established a new lineage in dogs (Hughes et al., 2012, Murcia et al., 2010, Hoelzer et al., 2010, Stack et al., 2013, Yamanaka et al., 2008). H3N8 canine influenza virus (CIV) originated from a direct transfer of a single H3N8 EIV to dogs within a large greyhound-training facility in Florida (Crawford et al., 2005). The virus, whose prototype strain is A/canine/Florida/2004, was then carried to several States of the USA by the infected greyhounds.
Other transfers of H3N8 EIV to dogs have also been reported (Daly et al., 2011, Murcia et al., 2011a), notably in 2002 in the UK (Daly et al., 2008), and in 2007 in Australia (Kirkland et al., 2010, Bryant et al., 2010) and Algeria (Myers C, 2006).
Of note, contemporary EIVs have also been found in pigs (Murcia et al., 2010), and camels (Yondon et al., 2013).

Despite close interactions between horses, dogs and humans, to date there has been no documented natural transmission of these viruses to humans (Khurelbaatar et al., 2014, Burnell et al., 2014). Direct experimental inoculation of human volunteers with EIV in the 1960s showed that infection did occur, although only to low levels (Kasel et al., 1965). This

suggests that high intrinsic barriers to the establishment of H3N8 viruses exist in humans, and onward transmission in two mammalian hosts (equine and canine) does not guarantee successful infection of another (human) (Parrish et al., 2015). Thus, EIV seems to represent a relatively low public health hazard.

1.1.7. A closer look at the virus

1.1.7.1. Genome and coding strategy

The genome of Influenza virus is approximately 13 kilobases and made of eight segments of single-stranded viral RNA (vRNA) of negative polarity. The vRNAs are packaged in distinct viral nucleoprotein complexes (vRNPs) held in a closed pseudo-circular conformation by the binding of the trimeric viral RNA-dependent RNA polymerase (PA, PB1, PB2) to the partially complementary 5' and 3' termini of the vRNA (Figure 1-1) (Eisfeld et al., 2015). The remainder of the vRNA segment is coated by viral NP oligomers arranged in an anti-parallel, superhelical, flexible rod-shape (Resa-Infante et al., 2011).

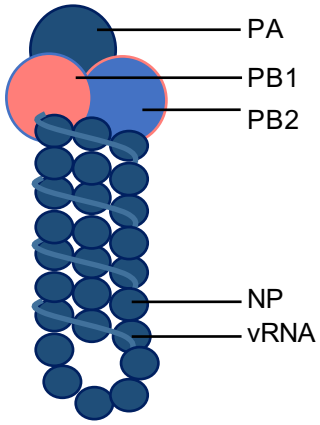


Figure 1-1: Schematic representation of an Influenza A virus vRNP

Influenza A virus ribonucleoprotein complex. The IAV genome is composed of eight ribonucleoprotein complexes (vRNPs), consisting of viral RNA (vRNA) encapsidated by viral ribonucleoproteins (NP) and a viral polymerase complex (PA, PB1, PB2) positioned at the extremity of the vRNA segment.

The IAV genome encodes up to sixteen proteins via alternative splicing or partially overlapping open reading frames (Stubbs and Te Velthuis, 2014, Szewczyk et al., 2014). The membrane-associated proteins HA and NA are encoded on segments four and six, respectively. The ion channel M2 and the matrix protein M1 are encoded on segment seven, either on a colinear mRNA for the latter or via alternative splicing for the former. The

constituent of the vRNP, including the three subunits of the viral polymerase (PB2, PB1, and PA) are encoded on segment one to three, respectively, and the NP protein on segment five. The non-structural proteins, NS1 and NEP (nuclear export protein, formerly known as NS2) are encoded on segment eight, and like M proteins, NS1 is encoded on a colinear mRNA, while NEP mRNA is generated via alternative splicing (Jagger et al., 2012). Other unspliced viral mRNAs encode PB1- and PA-related N-terminally truncated products, notably PB1-F2, PA-X. Finally, other viral proteins have been described (i.e. PB1-N40, PA-N155, PA-N182, M3, M4, M42, NS3 and NSP/NEG8), however, they are not uniformly expressed among strains and subtypes (Muramoto et al., 2013, Wise et al., 2009) and their existence in the case of EIVs remain uncertain.

1.1.7.2. Virion structure

Influenza A virions are pleomorphic and can be found both under a spherical (ranging from 80 to 120 nm) and filamentous form (up to 300 nm of diameter) (Elleman and Barclay, 2004, Harris et al., 2006, Vijayakrishnan et al., 2013, Elton et al., 2013). The biological reasons and consequences of this difference are still under investigation.

The viral envelope contains HA, NA and M2, which are linked via their cytoplasmic domains to an internal layer of M1. In turn, the latter interacts with the viral ribonucleoprotein (vRNP) (Hutchinson et al., 2012, Steinhauer and Skehel, 2002, Elton et al., 2013) NEP and NS1 have also been found in viral particles, while all other viral proteins have yet to be detected (Hutchinson et al., 2014) (Figure 1-2).

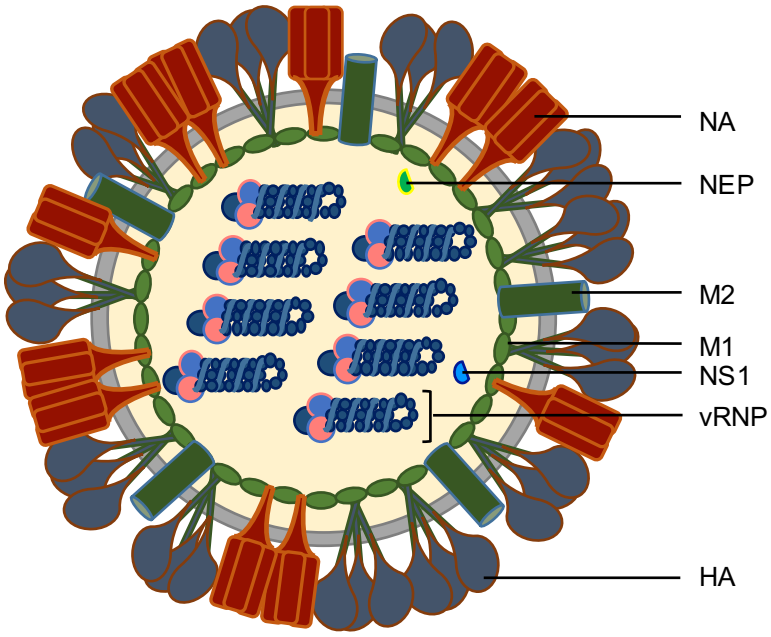


Figure 1-2: Schematic representation of an Influenza A virus particle

An IAV particle. Three viral proteins are embedded within the viral membrane: hemagglutinin (HA), neuraminidase (NA), and ion channel protein (M2). Matrix protein 1 (M1) underlies the viral envelope and holds the vRNPs inside the virion. NS1, a major virulence factor and NEP, involved in vRNP packaging into viral particles, are also present inside the viral particle.

1.1.7.3. Replication cycle

Once the NA protein has broken down mucus bonds within the respiratory mucus blanket, the virus gains access to the underlying epithelial cells. The HA protein can then bind sugar groups on the surface of epithelial cells (the primary target of IAVs), called sialic acids (Imai and Kawaoka, 2012). These sialic acids occur in a variety of modified forms in different hosts and tissues, which are likely to affect pathogen-host interactions and influence the outcome of infection (Wasik et al., 2017). After binding to the cell surface, the virus is internalized by endocytosis, and vRNPs are released in the cytoplasm before being imported into the nucleus where they serve as the templates for viral transcription and replication (Figure 1-3).

The vRNP is the minimum unit required for viral transcription and replication. The polymerase complex is responsible for initiating primary transcription, and it uses a unique method of transcription priming referred to as “cap-snatching” (Gu et al., 2015, Koppstein et al., 2015). The PB2 protein binds to the cap structure on the 5′-end of cellular pre-mRNAs, which is followed by cleavage of the pre-mRNA by the endonuclease function of PA. This produces a 5′-capped primer of ten to thirteen nucleotides, which is then used by the RNA-dependent RNA polymerase (PB1) to initiate transcription. Replication of the influenza virus

genome occurs through the synthesis of full-length complementary copies of the vRNA, the complementary RNA (cRNA), which are encapsidated into cRNPs and in turn act as templates for the polymerase-directed synthesis of progeny vRNPs. Early in infection, transcription is the predominant process allowing the synthesis of a pool of viral proteins. Subsequently, newly synthesised polymerase and NP are used to assemble and stabilise cRNPs or progeny vRNPs, which are generated through back replication of the anti-genomic cRNA intermediate. The same vRNA template can be either copied without modification into complementary RNA (cRNA) or transcribed into capped and poly-adenylated viral mRNA. (Reich et al., 2014). The viral polymerase is also responsible for poly-adenylating the 3'-end of viral transcripts, after which they are exported to the cytoplasm for protein translation in a similar fashion as host mRNAs (York and Fodor, 2013). vRNPs are selectively exported from the nucleus, transported to the plasma membrane (Lakdawala et al., 2014), and packaged together with other structural proteins into progeny virions, which bud from the cell surface (Rossman and Lamb, 2011). Specific RNA–RNA interactions between the different vRNPs are thought to ensure packaging of a full complement of genomic vRNPs (Giese et al., 2016, Hutchinson et al., 2010). Furthermore, M1 and M2 are thought to play a pivotal role in viral assembly and budding, while NA is important for the release of replicated virus from the host cell. (Steinhauer and Skehel, 2002)

Since viral replication and production of a new progeny is highly dependent on the cellular machinery, compatibility between the viral proteins and the molecular components and organelles of host cell is key in determining the outcome of an infection. This is particularly relevant for emergent viruses, which by definition are not adapted to replicate in the new host cells. Recently, two host factors have been described to specifically promote viral genome replication, ANP32A and ANP32B (also known as pp32 and APRIL, respectively) (Sugiyama et al., 2015), and the former has been identified as a species specificity host factor and shown to impact on the ability of avian IAVs to replicate in mammalian cells (Long et al., 2016).

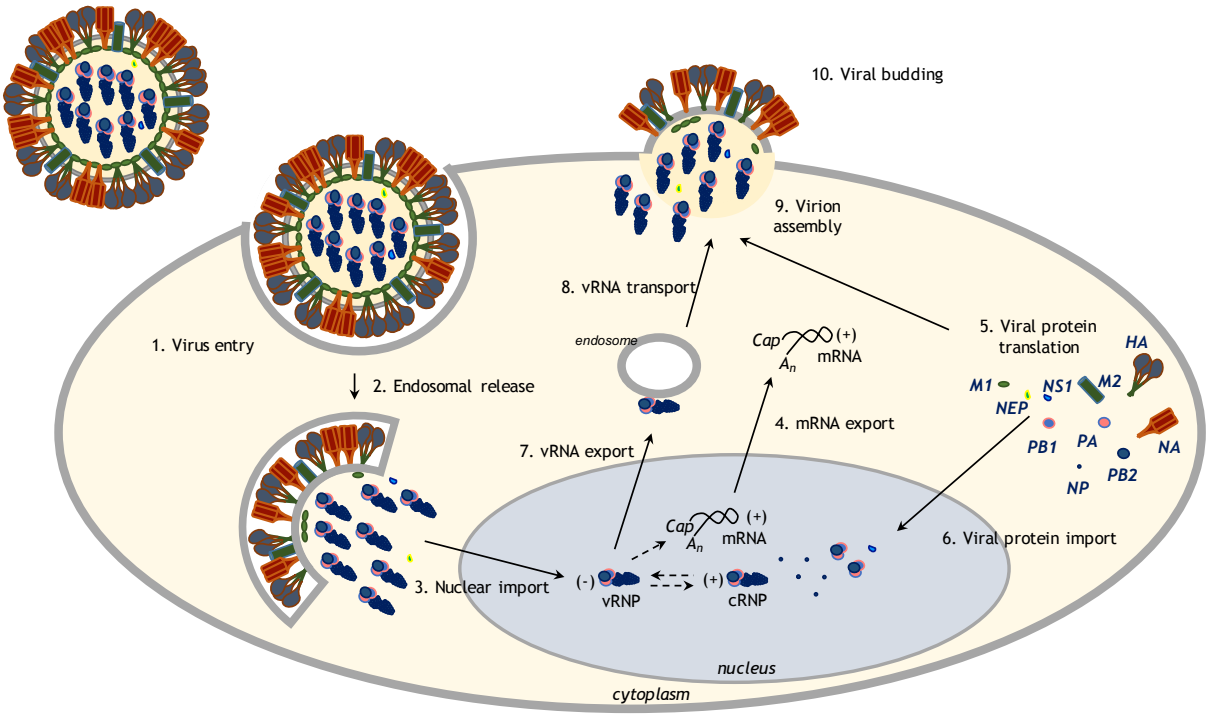


Figure 1-3: Influenza A virus replication cycle

Following binding of a virion to cell surface receptors containing sialic acid, the virion is internalised by endocytosis **(1)**. The endosomal membranes fuse with the viral membrane and release the vRNPs into the cytoplasm **(2)**. The vRNP are transported into the nucleus **(3)**, where the viral RNA polymerase transcribes the vRNA segments into mRNAs. These viral mRNAs are 5′ capped and 3′ polyadenylated before being exported into the cytoplasm **(4)** for protein translation by cellular machinery **(5)**. The viral RNA polymerase also synthetizes complementary RNA (cRNA) used as templates for the production of vRNA. The newly synthesised viral nucleoprotein and polymerase proteins are imported into the nucleus **(6)**, bind cRNA and vRNA and assemble vRNPs and cRNPs, respectively. Progeny vRNPs are exported out of the nucleus **(7)** and transported across the cytoplasm to the cell membrane **(8)**, where assembly of progeny virions takes place **(9)**. Mature virions incorporate a substantial amount of host proteins and are released by budding **(10)**.

Without lifelong protection against EIV and the threat of possible future pandemics, it is of great importance to have insight in how immunity against IAV infections is formed and how the virus manages to evade these immune responses.

1.2. Host defence against influenza A virus

In the healthy respiratory tract, the epithelium is ciliated and covered by a protective layer of mucus, which limits attachment of viral particles to the epithelium and entraps and transports them upward through the action of ciliary apparatus to the pharynx where they are swallowed. Once swallowed, viral particles are destroyed in the stomach by

the action of acidic secretions and digestive enzymes (Myers C, 2006). A normal architecture of the epithelium is essential for successful exclusion of viruses. However, if EIV breaks through this first line of defence and infects host cells, it will strongly compromise this epithelial barrier. Danger-associated molecular patterns (DAMPs) and pathogen-associated molecular patterns (PAMPs) will in turn be recognised by pathogen recognition receptors (PRRs) and alert the immune system that a viral invasion is occurring.

1.2.1. Detection of IAV

Influenza virus is recognized by at least 3 distinct classes of PRRs: Toll-like receptors (TLR), NOD-like receptor family member NOD-LRR- and pyrin domain-containing 3 (NLRP3), and cytosolic retinoic acid-inducible gene I (RIG-I)-like receptors (RLRs).

The Toll-like receptor family of transmembrane proteins are able to recognize a wide range of microbial ligands, and signal downstream through either myeloid differentiation primary response 88 (MyD88) or TIR-domain-containing adapter-inducing interferon-β (TRIFs) to stimulate expression of anti-viral cytokines, such as Interferons (IFNs) (see section 1.2.3 for more details about IFNs). Innate responses to IAV infection are primarily triggered by recognition of viral nucleic acids by endosomal TLR3 and TLR7. Their respective location enables the detection of viral nucleic acids present in the extracellular environment or produced by uncoating or degradation of viral particles (Wu et al., 2011). These features allow to mount an anti-viral response without the need for viral replication.

TLR3 has a relatively wide tissue distribution and is highly expressed in myeloid dendritic cells (mDCs) (Alexopoulou et al., 2001). Engagement of TLR3 by IAV dsRNA triggers a complex signal-transduction pathway starting with the dimerization and tyrosine phosphorylation of TLR3 (Sarkar et al., 2004), followed by the recruitment of TRIF (Yamamoto et al., 2003, Hoebe et al., 2003) and phosphatidylinositol 3 kinase (PI3K) (Sarkar et al., 2004). Engagement of TRIF activates both NF-κB and IRF-3 transcription factors, which in turn leads to IFN induction (Hardy et al., 2004).

TLR7 is expressed in few cell types and exclusively in endosomes (Kaminski et al., 2012). Ligands for this TLR include single-stranded RNA molecules (ssRNA) with no sequence specificity (Lund et al., 2004), but requiring the presence of several uridines in close proximity (Diebold et al., 2006). The mechanism of IFN induction follows a different profile from dsRNA-induced activation of TLR3, in that ssRNA-activated TLR7 recruits the MyD88 adaptor, which in turn recruits a complex containing the kinases interleukin-1 receptor-associated kinase 4 (IRAK-4), IRAK-1 and TRAF6 (Hacker et al., 2006), and leads to activation of NF-κB and induction of IFN. During an IAV infection TLR7 is particularly important in

plasmacytoid dendritic cells (pDCs), while it is dispensable in other cell types (Yoneyama et al., 2004, Jeisy-Scott et al., 2012), either because it is not expressed (Mayer et al., 2007) or because its localization is non-endosomal (Ioannidis et al., 2013).

NLRP3 forms a multiprotein inflammasome complex consisting of NLRP, pro-caspase 1, and the bipartite adaptor ASC (apoptosis-associated speck-like protein containing a carboxy-terminal CARD) (Lin et al., 2002). NLRP3 is expressed by myeloid cell types (monocytes, DCs, neutrophils and macrophages), as well as bronchial epithelial cells. The mechanism by which IAV activates the inflammasome is unclear, however, its importance has been proven using genetically modified mice lacking NLRP3. These mice have reduced viral-mediated inflammation in the lung, but increased mortality and viral clearance defects (Allen et al., 2009). Activation of inflammasomes results in the autocatalytic processing of pro-caspase 1 into its active form, which then cleaves pro-IL-1β and pro-IL-18 into IL-1β and IL-18, respectively (Wattrang et al., 2003). These activated cytokines then recruit monocytes and neutrophils into the lung, which are crucial to control infection and regulate tissue pathogenesis. In addition, the inflammasome participates in tolerance and tissue repair following an influenza virus infection (Iwasaki and Pillai, 2014). Furthermore, the role of NLRP3 in the protection against influenza virus is likely viral dose-dependent and maybe species-specific (Schaale et al., 2016).

Cytosolic RLRs are probably the most important group of PRRs during an IAV infection. This group of viral sensors consist of RNA helicase molecules and comprise RIG-I, Melanoma Differentiation-Associated protein 5 (MDA-5) and laboratory of genetics and physiology 2 (LGP2) (Pichlmair et al., 2006, Hornung et al., 2006).
RIG-I and MDA-5 recognize different types of RNAs absent in uninfected cells. RIG-I is most efficiently activated by short stretches of 5' triphosphate (5'-ppp) or 5' diphosphate (5'-pp) dsRNA, while MDA5 is activated by longer stretches of dsRNA in a 5' phosphate-independent manner (Pichlmair et al., 2006, Yoneyama et al., 2004, Hornung et al., 2006, Loo et al., 2008). In the case of IAV, viral detection and IFN production in airway epithelial cells, DCs, and alveolar macrophages is RIG-I-dependent (Le Goffic et al., 2007, Pichlmair et al., 2006, Van Reeth, 2000). MDA-5 may still contribute to IFN induction by influenza viruses in certain hosts, such as chickens that lack RIG-I (Husser et al., 2011, Liniger et al., 2012).

Both RIG-I and MDA-5 contain two caspase activation and recruitment domains (CARDs) at their N-terminus, as well as a RNA helicase domain possessing dsRNA-dependent ATPase activity, and a C-terminal regulatory domain (CTD). Upon ligand binding, an ATP-dependent

conformational change occurs in RIG-I, leading to exposition of its two CARDs (Myong et al., 2009, Kowalinski et al., 2011). These domains are then ubiquitinated by E3 ligases TRIM25 (Gack et al., 2009) and RIPLET (Oshiumi et al., 2013), and/or bound to free poly-ubiquitin chains produced by TRIM25 (Zeng et al., 2010). Once activated, RIG-I can then interact with a mitochondrial adaptor that acts as an intermediate between detection of viral RNA and downstream activation events. This adaptor was discovered in 2005 by four different groups and was referred to as mitochondrial antiviral signalling protein (MAVS) (Seth et al., 2005), virus-induced signalling adaptor (VISA) (Xu et al., 2005), CARD adaptor inducing IFN-β (Cardif) (Meylan et al., 2005), and IFN-β promoter stimulator protein 1 (IPS-1) (Kawai et al., 2005). This adaptor leads to the oligomerization and scaffolding of RIG-I into a multi-kinase signalling platform that will activate NF-κB and the IFN regulatory factors, IRF3 and/or IRF7, to stimulate expression of IFNs (Osterlund et al., 2007, Paz et al., 2006)(Figure 1-4-A).

LGP2 has a similar structure to RIG-I and MDA-5 but lacks the CARDs, and it has been proposed to function as a regulator of RIG-I and MDA-5 activity (Childs et al., 2013, Satoh et al., 2010) and IFN production (Rothenfusser et al., 2005). Indeed, LGP2 has been shown to potentiate IFN induction by synthetic mimic of viral dsRNA through co-operation with MDA-5. This co-operation is dependent upon dsRNA binding by LGP2, and the presence of helicase domain IV, both of which are required for LGP2 to interact with MDA-5. In contrast, LGP2 does not have the ability to enhance IFN induction by RIG-I, and instead acts as an inhibitor of RIG-I-dependent poly(I:C) signalling (Childs et al., 2013).

Interestingly, dendritic cells from mice harbouring a point mutation of Lysine 30 to Alanine in LGP2 (Lgp2 (K30A/K30A)), that abrogated the protein's ATPase activity, showed impaired IFN-β productions in response to various RNA viruses. Taken together, these data suggest that LGP2 facilitates viral RNA recognition by RIG-I and MDA5 through its ATPase domain (Satoh et al., 2010). Finally, the level of LGP2 expression seems to be also critical in determining the cellular sensitivity to induction by dsRNA, and this may be important for rapid activation of the IFN response at early times post-infection when the levels of inducer are low (Childs et al., 2013).

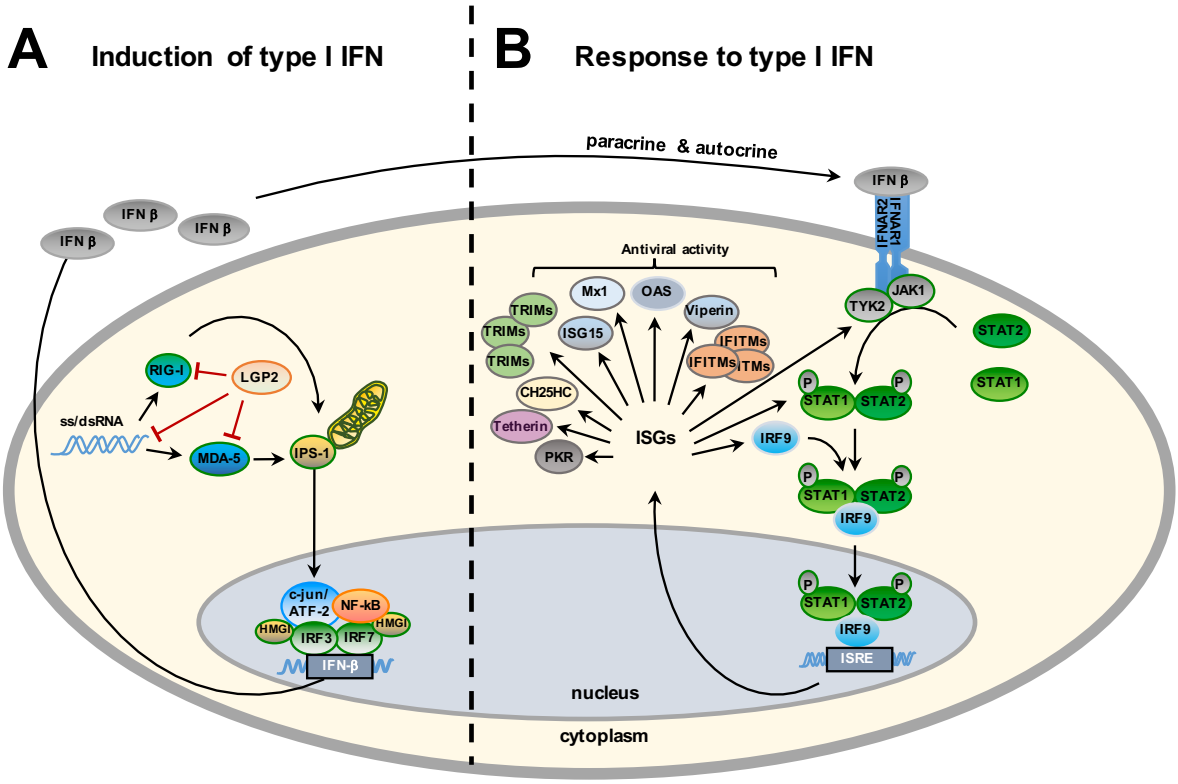


Figure 1-4: Type I IFN induction, signalling and action

(A) Induction of type I IFN. Single- and double-stranded (ss/ds)RNA, characteristic by-products of virus replication is recognised by RIG-I-like cytosolic sensors, RIG-I, MDA-5 and LGP2. Upon activation, RIG-I and MDA-5 will interact with a mitochondrial adaptor (IPS-1) and leads to activation of the transcription factors NF-κB, IRF-3/7 and c-jun/ATF-2 and formation of the enhanceasome. This complex recruits CBP and promote the formation of the basal transcriptional machinery and RNA polymerase II (not shown in this diagram), and allows full activation of the IFN-β promoter and production of IFN-β. Although not essential, formation of the enhanceasome is aided by HMGI. LGP2 has been proposed to function as a negative regulator of RIG-I and MDA-5 activity, and IFN-β induction. **(B)** Newly synthesized IFN-β binds to the type I IFN receptor (IFNAR) in a paracrine and/or autocrine manner. TYK2 then phosphorylates STAT2 (P-STAT2), while JAK1 phosphorylates STAT1 (P-STAT1). Following phosphorylation, P-STAT1 and P-STAT2 form a heterodimer that interacts with a monomer of Interferon Regulatory Factor 9 (IRF9). Together IRF9, P-STAT1 and P-STAT2 form the Interferon-stimulated gene factor 3 (ISGF3) complex. This complex is then transported into the nucleus, where it recognises a distinct DNA response element in the promoter of specific genes, the IFN-stimulated response element (ISRE). By binding to ISRE, the ISGF3 complex enhances transcription of hundreds of IFN-inducible or -stimulated genes (ISGs) that establish a cellular antiviral state (McBride et al. 2002). and activates the expression of numerous ISGs. Mx1, ISG15, OAS, PKR, TRIMs, IFITMs, CH25HC, tetherin, and viperin, among others are examples of induced proteins with antiviral activity. Furthermore, PRRs, IRFs, and several other signal transducing proteins, such as JAK, STAT1/2 and IRF9 are also induced by type I IFN. For details see text.

1.2.2. Natural Influenza PAMPs

It is likely that distinct PAMPs are generated at different stages of the virus life cycle. Several studies have indicated that incoming influenza A vRNPs are not sufficient to induce IFN during infection of epithelial cells and monocyte-derived DCs (moDCs), and that viral

RNA synthesis is required (Crotta et al., 2013; Killip et al., 2014; Osterlund et al., 2012). Indeed, efficient inhibition of IRF3 and NF-kB activation by the cellular transcription inhibitors actinomycin D and alpha-amanitin in influenza-infected cells strongly suggests that incoming genomes do not function as a major PAMP in lung epithelial cells (Killip et al., 2014). Furthermore, expression of IFNs following influenza A virus infection correlates with the accumulation of viral RNAs (Osterlund et al., 2012) and can be completely abrogated following inactivation of the virus by heat or UV treatment (Crotta et al., 2013; Osterlund et al., 2012). A minority of incoming IAV genomes may contribute to IFN induction at very early times post-infection, it appears that the viral polymerase products generated later in infection represent the most significant IAV PAMPs (Osterlund et al., 2012, Crotta et al., 2013). Interestingly, the above studies have been limited to a relatively small number of influenza A virus strains, so it is conceivable that different influenza A virus strains may vary in their capacity to be recognised by RIG-I during vRNP entry.

Taken together, these observations strongly suggest that the generation of the major influenza A virus PAMPs requires the synthesis and nuclear export of an RNA product or products from incoming genomes, and that these RNAs can be generated even in conditions where cRNP and vRNP accumulation is impaired. Nevertheless, it is likely that distinct PAMPs are generated at different stages of the virus life cycle; thus, a minority of incoming influenza A virus genomes may contribute to IFN induction at very early points post-infection, with viral polymerase products (including, but not limited to, progeny genomes) functioning as a more significant PAMP population later in infection (Killip et al., 2015).

Evidence also suggests that activation of the IFN pathway upon infection is not associated with ‘normal’ virus replication. The link between the presence of ‘abnormal’ defective interfering (DI) particles and IFN induction by negative-strand RNA viruses is well known. DI particles, first identified in the 1950s as ‘incomplete’ non-infectious viruses (Von Magnus, 1951) are readily generated during an IAV infection. They appear to be limited to internal deletion of RNA segments that retain their 5’ and 3’ termini. Influenza virus DIs can arise from all genome segments, but are more frequent from PB2, PB1, PA segments.

These DI particles are distinguished from non-defective particles by their inability to complete a full replication cycle due to the absence of the coding region for one or more viral factors otherwise essential for replication. DI viruses are therefore only able to replicate when a co-infecting ‘normal’ virus is present, hence providing the missing factors. Their relevance in vivo has been questioned for a long time, however, relatively recent studies have confirmed their role in natural infections, and internal deletion DI viruses have been detected in samples from infected mice (Tapia et al., 2013), poultry (Jonges et al., 2014) and humans (Saira et al., 2013, Jonges et al., 2014). DIs are generated de novo during

in vivo infections (Tapia et al., 2013) and the transmission between patients has been described (Saira et al., 2013). Studies have also correlated the IFN-inducing phenotype of influenza viruses with a propensity to generate or accumulate high numbers of DI genomes in tissue culture (Perez-Cidoncha et al., 2014). What remains unclear is why DI RNAs should be more effective at inducing IFN than non-defective full-length genome segments since they are encapsidated into stable nucleocapsids and should therefore be hidden from recognition by viral sensors (Strahle et al., 2006). Their ability to induce IFN may simply be due to their faster replication rate compared to full-length genomes because of their smaller size. Alternatively, DI-mediated interference with viral polymerase and NP expression may lead to a reduction in polymerase-induced shut-off of cellular gene expression (Ngunjiri et al., 2012) or affect the efficiency of cRNA and vRNA encapsidation. Furthermore, DI-mediated interference may reduce NS1 expression, thereby contributing to IFN induction by limiting its antagonism by the virus. Perhaps also, smaller segments of genomic RNAs form a panhandle more readily than full-length RNA when they are unencapsidated, or DI vRNPs may be less stable than full-length vRNPs and thus more prone to release free RNA. Finally, very small influenza genome templates can be replicated in the absence of NP (Resa-Infante et al., 2010). Thus, it is a possibility that small DI RNAs could be replicated without being encapsidated and these products could activate more strongly IFN induction.

1.2.3. Interferons

Interferons (IFNs) represent a group of secreted cytokines that elicit distinct antiviral effects. There are divided in three classes according to their amino acid sequence and the type of receptor through which they signal.

The type I IFNs, discovered in 1957 (Isaacs and Lindenmann, 1957) comprises a large group of molecules, including multiple IFN-α genes, up to three IFN-β genes, and IFN-ω, -ε, -τ, -δ, -κ genes. IFN-α and -β are induced directly in response to viral infection, whereas IFN-ω, -ε, -τ, -δ, -κ are thought to have different functions. Indeed, IFN-ε is constitutively expressed by epithelial cells of the female reproductive tract and is hormonally regulated. IFN-ε-deficient mice have notably an increased susceptibility to infection of reproductive tract by the common sexually transmitted infections herpes simplex virus 2 and Chlamydia muridarum. Thus, IFN-ε is believed to be a potent anti-pathogen and immune-regulatory cytokine (Fung et al. Science 2013).

IFN-τ expression is dependent of the presence of a trophoblast. Polymorphisms have been identified in cattle, sheep and goats, and genes and gene alleles encoded proteins do not display identical antiviral, antiproliferative and antiluteolytic activities. The need for distinct IFN-τ activities remains debatable, but the consensus is that this complexity in expression and biological activity must be needed to provide the best opportunity for pregnancy to be recognized by the maternal system so that gestation may continue (Ealy et al., 2017).
Finally, IFN-κ is constitutively expressed in keratinocytes and acts as a type I IFN in hr-HPV-positive keratinocytes, inducing ISG expression following viral infection (Habiger et al., 2015).

In horses, thirty-two putative type I IFN loci on chromosome twenty-three have been identified. A phylogenetic analysis characterized them into eight different type I IFN classes, including six IFN-α, four IFN-β, eight IFN-ω (plus four pseudogenes), three IFN-δ (plus one pseudogene), one IFN-κ, one IFN-ε, and 3 loci belonging to what has previously been called IFN-αω. Results from cell cultures also showed that leukocytes readily expressed IFN-α, IFN-β, IFN-δ, IFN-μ, and IFN-ω mRNA on induction with live virus; while fibroblasts only expressed IFN-β mRNA on stimulation (Detournay et al., 2013).

From this point onwards, the term 'type I IFN' will be used to refer to the virally induced IFN-α and -β.

Optimal induction of type I IFN requires the formation of an enhanceasome formed by IRF-3/IRF-7, NF-κB and c-jun/ATF-2 (Figure 1-4-A). This complex participates in the recruitment of CREB-binding protein (CBP)/p300, which in turn promotes the assembly of the basal transcriptional machinery and RNA polymerase II. Each of these transcription factors can bind the IFN-β promoter with limited affinity, and optimal induction requires cooperativity between these factors. Only IRF-3 and IRF-7 are indispensable for IFN-β induction, while activation of both NF-κB and c-jun/ATF-2 may not be essential (King and Goodbourn, 1994, Ellis and Goodbourn, 1994, Goodbourn et al., 1985). Formation of the enhanceasome can be aided by the high-mobility group (HMG) chromatin-associated protein I (HMGI, also known as HMGA) (Munshi et al., 2001, Merika and Thanos, 2001), although this protein is not absolutely essential (Berkowitz et al., 2002).
Interestingly, type I IFN itself can enhance the its yield in a process known as 'priming'. It is believed that direct induction of primary type I IFN genes (IFN-β and α4 in mouse) (Erlandsson et al., 1998, Marie et al., 1998) takes place via IRF-3. Then, by positive

feedback, IFN-β induces the synthesis of IRF-7, which in the presence of a continued infection, enhances dramatically the transcription of primary IFN genes and allows the transcription of secondary IFN genes (other IFN-α genes) (Marie et al., 1998, Prakash et al., 2005).

All type I IFNs bind a heterodimeric receptor, Interferon-α/-β receptor (IFNAR) composed of two chains (IFNAR-I and -II). Prior to activation, the cytoplasmic tail of IFNAR-I and -II is associated with Tyrosine Kinase 2 (Tyk2) and Janus Kinase 1 (JAK1), respectively (Figure 1-4-B). Signal transducer and activator of transcription 2 (STAT2) is also bound to IFNAR-II before induction and is weakly associated with STAT1 (Precious et al., 2005). When the type I IFNs bind the IFNAR receptor, a conformational change occurs allowing phosphorylation of IFNAR-I by Tyk2 on tyrosine 466, which creates a strong docking site for STAT2. Tyk2 then phosphorylates STAT2 (P-STAT2) on tyrosine 690, while JAK1 phosphorylates STAT1 (P-STAT1) on tyrosine 701. Following phosphorylation, P-STAT1 and P-STAT2 form a heterodimer, which creates a novel nuclear localisation sequence (NLS) (Banninger and Reich, 2004) and inactivates the dominant constitutive nuclear export of STAT2 (Frahm et al., 2006). In addition, the transcriptional co-factor CREB-binding protein (CBP) is recruited to IFNAR-II and catalyses its acetylation, which creates a docking site for a monomer of Interferon Regulatory Factor 9 (IRF9). Together IRF9, P-STAT1 and P-STAT2 form the IFN-stimulated gene factor 3 (ISGF3) complex. Of note, acetylation of IRF-9, P-STAT1 and P-STAT2 may also aid in the assembly of the ISGF3 complex (Tang et al., 2007). The ISGF3 complex is then transported into the nucleus, where it recognises a distinct DNA response element (sequence 5'-AGTTTNNNTTTCC-3') in the promoter of specific genes, the IFN-stimulated response element (ISRE). By binding to ISRE, the ISGF3 complex enhances transcription of hundreds of IFN-inducible or -stimulated genes (ISGs) that establish a cellular antiviral state (McBride et al., 2002).

The remaining two classes of IFN are represented by type II and III IFNs. Type II IFN has a single member, IFN-γ, and is secreted by activated T cells and natural killer (NK) cells, rather than in direct response to viral infection. Hence it will not be discussed further in this review. In contrast, type III IFNs have been described more recently and comprise IFN-λ1, λ2 and λ3, also referred to as IL-29, IL-28A and IL-28B, respectively (Randall and Goodbourn, 2008). These cytokines appear to use the same pathways as type I IFN to sense viral infection, and they are directly upregulated following virus infection. Furthermore, the panel of ISGs upregulated in airway epithelia in response to type I and type III IFNs following an IAV infection completely overlaps (Crotta et al., 2013). Several studies suggested that airway epithelial cells are also primarily responsible for type III IFN

production, while plasmacytoid dendritic cells and alveolar macrophages are the major source of type I IFNs (Crotta et al., 2013, Cheung et al., 2002, Ioannidis et al., 2013). Type III IFNs signal through a heterodimeric cell surface receptor composed of the β subunit of the IL10-receptor and IFN-λ1 receptor. Unlike the type I IFN receptor, the type III IFN receptor shows a limited tissue distribution (Crotta et al., 2013, Mennechet and Uze, 2006, Zhou et al., 2007), and is primarily expressed on epithelial cells of respiratory and gastrointestinal tracts. The role of type III IFNs during an IAV infection remains to be fully established, however, they are believed to elicit an equivalent antiviral response to type I IFN in airway epithelia, while restricting immune cell activation (Crotta et al., 2013, Onoguchi et al., 2007).

The IFN response to viral infection is tightly controlled, and shortly after IFN exposure cultured cells enter an IFN-desensitized state that can last for several days. This desensitized state is supposed to allow cells to recover from IFN signalling. This state is established in cells by multiple mechanisms, either intrinsic or mediated by the actions of ISGs. Suppressor of cytokine signalling (SOCS) proteins are induced early in the IFN response and play an important role in early IFN desensitization (Schneider et al., 2014). These proteins inhibit JAK-STAT signalling by binding to phosphorylated tyrosine residues, on either the IFN receptor or the JAK proteins, resulting in inhibition of STAT binding as well as JAK activity. Similarly, in the absence of IFN signalling an elevated baseline level of SOCS1 protein expression renders cells less responsive to IFN. Receptor endocytosis and turnover play also an important role in reducing the level of JAK-STAT signalling. IFN signalling is further decreased by the action of phosphatases that inactivate the JAKs and STATs. In addition, STAT activity can be modulated by members of the family of protein inhibitors of activated STAT (PIAS) (Schneider et al., 2014), however the precise mechanisms by which these proteins inhibit IFN signalling remain unclear.

1.2.4. Interferon-stimulated genes

ISGs are genes whose transcriptional output increases in response to virus-induced IFN, largely due to the presence of ISRE sequences in promoter and enhancer regions. Their function is to establish a cellular anti-viral state to efficiently limit further replication and spread of the virus (Randall and Goodbourn, 2008). ISGs can represent up to 5% of the total number of cellular genes, however, this percentage may be underestimated as it only represents genes with changes in mRNA abundance. Since several ISGs are present in

multiple isoforms and give rise to proteins with differences in cellular localization and activity, some ISGs changes upon type I IFN stimulation may remain undetected.

ISGs take on a wide range of activities. Many will control viral infection by directly targeting pathways and functions required during pathogen life cycles. Indeed, to complete their life cycle, viruses must enter cells, replicate their genome, and exit in order to infect new cells. Every stage of the virus life cycle is a potential target for ISGs intervention. In addition, a number of ISGs also act to enhance pathogen detection and innate immune signalling, while others encode for pro-apoptotic proteins leading to cell death under specific conditions. Of these ISGs, several have been identified as having direct anti-IAV activities.

Among important and well-studied proteins encoded by ISGs, the myxovirus resistance proteins (Mx1 and Mx2) are considered important in the defence against IAV infection. Furthermore, diversity of equine Mx gene has been associated with variations in susceptibility vis-a-vis resistance against EIV evidence, which suggests an important role of these proteins for the resistance to influenza virus infection (Manuja et al., 2014). Mx1 and Mx2 belong to a small family of dynamin-like large guanosine triphosphates (GTPases). The former broadly inhibits IAV and acts prior to genome replication at a post-entry step in the virus life cycle. It contains a middle stalk domain and a GTPase effector domain, which are both essential for self-oligomerization and formation of ring-like structures that trap incoming viral components, such as NP. In doing so, Mx1 inhibits nuclear import of vRNP, and potentially directs them to sites of degradation. Mx2 has been characterized more recently as an antiretroviral effector protein against HIV but has less potent or no antiviral activity against IAV (Busnadiego et al., 2014).

Many of the proteins involved in ISGylation and de-ISGylation are also induced by type I IFN (i.e. TRIM25, ISG25, USP18, …) (Sgorbissa and Brancolini, 2012). ISGylation can positively or negatively affect the targeted protein. In the case of IRF3 it increases stability by preventing poly-ubiquitination and leads to sustained transcription factor activity. In contrast, ISGylation of cyclin-dependent kinase 1 (CDK1) leads to protein destabilization and cell cycle inhibition. It is believed that ISGylation is a general, non-specific mechanism of host defence, which potentially impacts all viral proteins translated in IFN-stimulated cells (Durfee and Huibregtse, 2010). A small IFN-induced ubiquitin-like protein, Interferon-stimulated gene 15 (ISG15) is one of the most highly induced ISGs. However, despite reports describing ISG15 function and targets in vitro, its role in vivo remains controversial

(Bogunovic et al., 2012). ISG15 is also secreted from immune cells and enhances the production of type I IFN (Bogunovic et al., 2013). Removal of ISG15 conjugates (deISGylation) is performed by USP18. USP18 maintains long-term IFN desensitization through an interaction with IFNAR-II, which suggests that USP18-mediated inhibition may be restricted to type I IFN signalling (Schneider et al., 2014).

The tripartite motif (TRIM) family of proteins are also induced by virus-induced IFNs. This family is composed of more than 60 members and exhibits a wide range of activities including E3 ubiquitin ligase activity, SUMOylation and ISGylation. Among them is found TRIM25 responsible for ubiquitination and activation of RIG-I as described previously (section 1.2.1).

Another important ISG-encoded protein is the 2'-5'-oligoadenylate synthase (OAS). Upon binding dsRNA, OAS becomes enzymatically active and converts ATP into 2'-5'-oligoadenylate, which then functions as a second messenger to activate latent RNase L. OAS and activated RNAse L act together to degrade viral RNA in the cytosol. They also generate short fragments with 3' monophosphates, which lead to protein synthesis inhibition and viral growth arrest. By-products of this degradation might also be recognized by RLRs, hence participating in the positive feedback in the IFN response (Silverman et al., 2007). In accordance with this, the magnitude of IFN induction is enhanced by the activation of RNase L (Malathi et al., 2007).

The double-stranded RNA (dsRNA)-activated protein kinase (PKR) is an RNA-binding protein kinase constitutively expressed in an inactive conformation in mammalian cells. The binding of dsRNA to PKR release its auto-inhibition and thereby activates it. PKR also becomes upregulated upon interferon treatment (Hovanessian, 1989). Once activated, it phosphorylates a number of substrates and activates different signal transduction pathways to counteract potential threats. PKR notably phosphorylates the α-subunit of eukaryotic translation initiation factor 2α (EIF2α) in parallel of binding viral dsRNA, thus inhibiting translation of both cellular and viral proteins (Garcia et al., 2006). Moreover, PKR plays an important role in signal transduction and transcriptional control through activation of IκB, the inhibitor of NF-κB. Activation of PKR in infected cells results in apoptosis, cell growth arrest and autophagy, all of which limit viral replication and spread in the host. In addition, PKR has been shown to stabilize IFN-α/β mRNAs, thereby ensuring robust type I IFN production (Munir and Berg, 2013).

The Cholesterol 25-Hydroxylase (CH25HC) gene is also upregulated by type I IFNs. The product of this gene is an enzyme that converts cholesterol into 25-hydroxycholesterol (25HC). CH25H mediates protection against IAV at an early stage in the infectious cycle, possibly at the step of virus-host membrane fusion (Schoggins and Randall, 2013). However, 25HC may impact virus infection by additional mechanisms, such as directly altering membrane properties, inhibiting sterol biosynthesis, and affecting prenylation of both virus and host proteins, which plays an important role in the virus life cycle (Matsumiya et al., 2013). In addition, high concentrations of 25HC change the physical properties of membranes, which prevents virus-host membrane fusion (Matsumiya et al., 2013).

The only ISGs shown to have a bona fide role in blocking virus entry are members of the IFN-inducible transmembrane (IFITM) family. The IFITM family of proteins is composed of four members, IFITM1, IFITM2, IFITM3 and IFITM5, that have been shown to be potent inhibitors of IAV infection (Brass et al., 2009). These proteins are enriched in late endosomes and lysosomes. The most potent member of the IFITM family against IAVs is IFITM3, which interfere with fusion between viral and endosomal membranes thereby limiting the viral entry (Brass et al., 2009, Everitt et al, 2012).

During late stages of the virus life cycle, viral nucleic acids are packaged into capsids, and viral particles exit cells either by cell lysis, exocytosis, or direct budding from the plasma membrane. Relative to other stages in the virus life cycle, few ISGs are known to inhibit viral assembly and viral egress. Viperin, encoded by the ISG RSAD2, is one of the best-studied, most highly induced antiviral effectors. It can be induced by at least two different innate immune pathways, such as the JAK-STAT signalling pathway or via direct activation by IRF3. Viperin normally resides in the endoplasmic reticulum (ER) and in ER-derived lipid droplets, organelles important for lipid metabolism. Viperin inhibits IAV budding at the host cell membrane by inhibiting farnesyl diphosphate synthase (FFPS), an enzyme involved in isoprenoid biosynthesis (Wang et al., 2007).

Another ISG-encoded protein important at the late stage of the virus life cycle is Tetherin. This protein is encoded by the ISG BST2, and it inhibits virus budding by using two membrane anchors to trap virions on the plasma membrane (Swiecki et al., 2011).

Finally, PRRs, IRFs, and several other signal transducing proteins, such as JAK, STAT1/2 and IRF9, are present at baseline, but are also ISGs and reinforce the IFN response. Upon induction, this set of ISGs are directly induced by IRF activation independently of the

JAK-STAT pathway, which likely evolved to counteract pathogen strategies of immune evasion (Iwasaki and Pillai, 2014).

1.2.5. Other immune players

Although innate immune responses control virus replication during the early phase of influenza virus infection, clearance and recovery from infection requires adaptive immune responses (van de Sandt et al., 2012). Virally infected epithelial cells become the targets of natural killer cells (NK cells), which are cytotoxic lymphocytes of the innate immune system responsible for viral clearance. In addition, monocytes and neutrophils are rapidly recruited to the lung after an IAV infection, and together with alveolar macrophages they help clear dead cells. In addition to alveolar macrophages, tissue-resident macrophages are found within the pulmonary interstitia. These cells have the capacity to transport antigens from the lungs to the draining lymph nodes and they may serve as antigen presenting cells (APCs) for activation of effector T cells (Iwasaki and Pillai, 2014). However, the professional APCs are conventional dendritic cells (cDCs). They reside below the airway epithelial cells within the pulmonary interstitia, and extend processes between airway epithelial cells (Iwasaki and Pillai, 2014). They are crucial for the induction of adaptive immune responses by the presentation of viral antigens to naïve and memory T and B lymphocytes. Plasmacytoid DCs (pDCs), distinct from cDCs, are activated during influenza virus infection. These cells are weak activators of naïve T cells but are major producers of type I IFNs.

B lymphocytes are the first cells of the adaptive immune system to encounter the virus (Iwasaki and Pillai, 2014). Once activated, they differentiate into plasmocytes, also known as antibody-producing cells. In the lung, these cells generate mainly IgG and IgM, whereas in the upper respiratory tract the vast majority of antibodies produced are IgA (Waffarn and Baumgarth, 2011). Antibodies are generated against most viral proteins, although at different levels and kinetics. Best understood are the strong and often neutralizing responses against HA. Serum antibody titres are detected around six to seven days post-infection and are delayed by at least three days compared with the respiratory tract. Antibody titres increase for about one month, and IgG and IgA memory B cells are maintained for months in the respiratory tract (Waffarn and Baumgarth, 2011). Virus-specific antibody-producing cells also reside transiently in the spleen, starting around six to seven days post infection, but remain in the bone marrow in the long-term. Interestingly, a small subset of B cells, distinct from conventional B cells, contribute to protection from IAV

infection by generating natural IgM, even prior to any encounter with the virus. They contribute to host protection by virus neutralization and are required for maximal induction of IAV-specific IgG. Moreover, they are strong producers of IL-10, and could be involved in the regulation of local immune responses (Braciale et al., 2012).

T lymphocytes can be categorized into three groups: T cells responsible for lysis of infected cells by exocytosis of perforin- and granzyme-containing granules; T cells that can induce apoptosis of infected cells by expressing CD95 ligand (also known as FAS ligand) or TNF-related apoptosis-inducing ligand (TRAIL); and T cells that can produce pro-inflammatory mediators and regulatory mediators in response to encounter with virus-infected cells (van de Sandt et al., 2012).
Recent evidence strongly suggests that activation and function of effector T cells in virus-infected lungs are sculpted by the local milieu generated by the resident and infiltrating inflammatory cells, as well as by factors produced in response to virus infection (Braciale et al., 2012).

1.2.6. Particularities of the athletic horse

Exercise is considered a physical stress and has been shown to have profound effects on the immune system of man, pigs, dogs, cattle and rats (Tiollier et al., 2005, Pedersen and Toft, 2000). In horses, evidence suggest that a strenuous exercise is associated with a reduction of TNF-α and IFN-β production by pulmonary leukocytes treated with poly(rI).poly(rC) (structurally similar to double-stranded RNA) for 24h compared to controls (Folsom et al., 2001). In addition, after five days of intense exercise a significant decrease of lympho-proliferation was noted as well as a decrease in type II IFN expression, which is mainly produced by CD4+ helper T cells (of Th1 subtype), CD8+ cytotoxic T cells (Mignot et al., 2012). Taken together, these data suggest that young untrained horses subjected to strenuous exercise have an increased susceptibility to EIV infection.

1.3. Virus strategies to counteract type I IFN

When IAV enters the body, a race begins between the host immune system and the infecting virus. Eventually, most viral infection will be successfully controlled by the host immune response, however, IAVs have evolved various strategies to control the early step

of this response, by notably blocking the type I IFN system, in order to open a window for replication and transmission and thus survive in nature.

1.3.1. Non-structural protein 1

NS1 is the main viral antagonist of the host interferon response in infected cells. It can block these responses at multiple levels with a striking degree of redundancy. More importantly, viruses unable to express NS1 are only able to replicate in type I IFN compromised systems (Garcia-Sastre et al., 1998, Kochs et al., 2007c), which demonstrates how relevant for the virus is to possess an NS1 protein blocking efficiently the host IFN responses.

As previously mentioned, NS1 is encoded on a colinear mRNA derived from the NS segment, and upon splicing, this segment also encodes NEP mRNA (Inglis et al., 1979, Lamb and Choppin, 1979) (Figure 1-4). NS1 and NEP share their first 10 amino acids (nucleotide 27 to 57 on the NS segment), while nucleotides 530 to 719 are shared between NS1 and NEP ORFs, however on a shifted reading frame (Lamb and Lai, 1980).

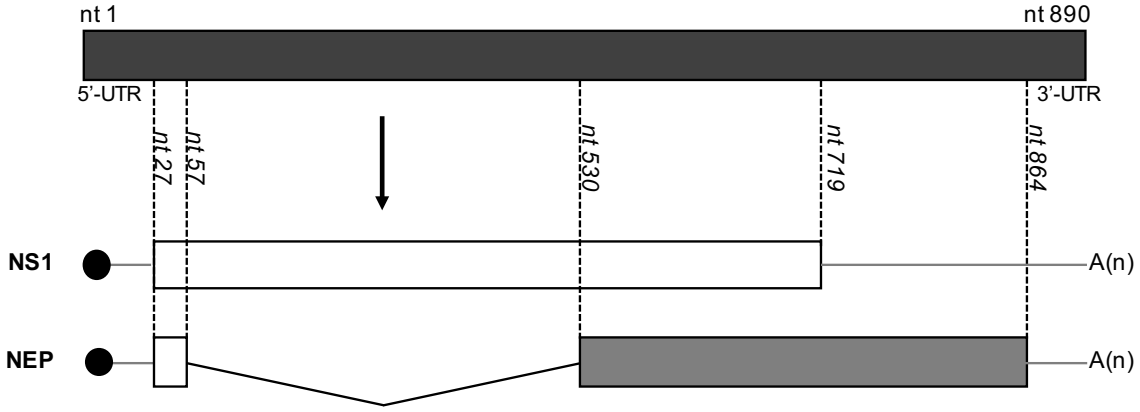


Figure 1-5: NS segment coding strategy.

Schematic representation of the NS genes and gene transcripts. In-frame NS1 mRNA is indicated by a white box, with nucleotide length indicated by vertical dotted lines and numbers. Viral NEP mRNA is also shown, with white box indicating the in-frame mRNA sequence shared between viral NS1 and NEP ORFs, and grey box representing the unique ORFs of the viral NEP mRNA transcript. 5'-cap and 3'-polyA tails are indicated by a black circle and A(n) joint to each ORF, respectively. 5'- and 3'-UTRs are indicated. UTR: untranslated region, A(n): polyA tail, nt: nucleotide, ORF: Open-reading frame.

In infected cells, the steady-state amount of spliced NEP mRNA is approximately 10% of un-spliced NS1 mRNA (Lamb and Lai, 1980). Regulation of splicing is controlled in part by NS1 itself and may represent a mechanism of autoregulation of protein levels within infected cells (Garaigorta and Ortin, 2007).

NS1 mRNA translates into a 26 kDa polypeptide, which varies in length. EIV NS1 ranges between 217 and 230 amino acids, and in IAVs isolated from other host species NS1 can be as short as 202 amino acids and as long as 237 (Dundon and Capua, 2009). Furthermore, NS1 proteins can be divided into three distinct variants: alleles A and B, and H17-H18 bat virus NS1 proteins (Ludwig et al., 1991, Turkington et al., 2015, Scholtissek and von Hoyningen-Huene, 1980). Among the same allele, 90 to 100% of residues are conserved, while between alleles conservation can be as low as 60%. As for the majority of mammalian IAVs, all EIVs possess an allele A NS1 protein. Allele B NS1s are found in avian IAVs (Guo et al., 1992), and may represent an archaic version of NS1, which has undergone strong selection pressure and mutated to give rise to allele A after entering in mammalian host populations (Ludwig et al., 1991).

1.3.1.1. Structure

The NS1 polypeptide comprises a N-terminal RNA binding domain (RBD) (amino-acids 1-73) formed of three α-helices, and a C-terminal effector domain (ED) (residue 85 to end) made of seven β-strands and three α-helices, separated by a short linker region (Qian et al., 1994) (Figure 1-5). NS1 RBD can homodimerize in a six-helix antiparallel bundle and form a tract of conserved basic and hydrophilic residues capable of binding with low affinity to several RNA species. The binding is independent of the RNA sequence and follows a dimer:dsRNA stoichiometry of 1:1 (Chien et al., 2004, Hatada and Fukuda, 1992, Qian et al., 1994). Residue R38 and K41 are particularly important for this function, and additional amino acids also contribute (i.e. T5, P31, D34, R35, G45, R46, T49) although to a lesser extent (Liu et al., 1997, Chien et al., 1997). NS1 ED can also dimerize and W187, located at the helix-helix interface, is essential for it (Hale et al., 2008a). It has been proposed that NS1 could adopt various dimeric states in a strain- or ligand-specific manner, which would likely contribute to NS1 multifunctional potential. The short linker region between RBD and ED is thought to be flexible and would allow NS1 homodimer to adopt different conformations. Additionally, the linker length and amino acid composition, as well as the nature of neighbouring residues are likely to be involved in NS1 quaternary structure. Evidence also suggests that the RBD association is highly stable and has the properties of an

obligate dimer, whereas the ED dimerization is likely to be transient (Kerry et al., 2011a, Aramini et al., 2011). Furthermore, it is not clear whether ED dimerization takes place within the same RBD dimer or with neighbouring dimers to promote NS1 multimerization, or both (Wang et al., 2002, Bornholdt and Prasad, 2006, Carrillo et al., 2014). Importantly, it has been suggested that the tridimensional disposition that NS1 can adopt may influence its interactions with cellular partners in a strain specific manner (Ayllon and Garcia-Sastre, 2015).

There is currently no structure available for the last 25 amino acids of NS1. This part of the protein is thought to form an intrinsically disordered and flexible "tail", which may only adopt an ordered structure upon binding to the appropriate ligand (Hale et al., 2008c). Interestingly, this is one of the most variable regions of the protein and the main cause of diversity in NS1 lengths. Indeed, naturally occurring C-terminal truncations are relatively common among IAVs (Suarez and Perdue, 1998, Hale et al., 2010, Dundon and Capua, 2009). For example, EIV NS1 was initially composed of 230 amino acids, however, in the late 1990's a non-sense mutation of codon 220 resulted in a loss of eleven amino acids at its C-terminus (Rivailler et al., 2010, Dundon and Capua, 2009, Barba and Daly, 2016). Of note, three isolates from the early 1970's (A/equine/Sachiyama/1971; A/equine/Tokyo/2/1971; A/equine/Algiers/1/1972) harboured a 217-amino-acid-long NS1 protein. Work done by Quinlivan and colleagues on EIV NS1 deletion mutants revealed that the length of the protein did not correlate with the level of viral attenuation as initially thought, and they speculated that the basis for EIV attenuation was related to the level of expression of the protein rather than its length (Quinlivan et al., 2005)).

Variation in NS1 length has also been seen in human isolates, which acquired seven additional residues in their C-terminus in the late 1940s, extending NS1 length from 230 to 237 amino acids (Hale et al., 2008c). For the following 40 years, all NS1 proteins of human isolates (H1N1, H2N2 and H3N2) were 237 residues long. This extension was retained until the late 1980s, when both co-circulating H1N1 and H3N2 viruses reverted to 230 amino acids. Despite some evidence of being functionally implicated in NS1 nuclear and nucleolar localization, there is no clear evolutionary advantage associated with these extension or truncation events (Melen et al., 2007).

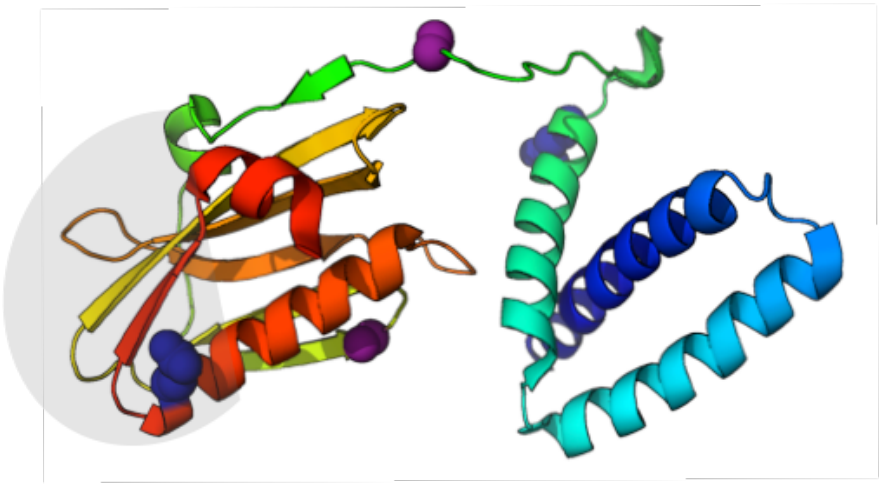


Figure 1-6: Tri-dimensional structure of NS1.

Cartoon representation of NS1. The full length NS1 protein of H6N6 avian virus has been used for this cartoon representation (Carrillo et al., 2014, Bornholdt and Prasad, 2006). Of note, to solve the structure of the full length NS1, the introduction of a double substitution R38A-K41A was necessary to prevent aggregation of the recombinant protein. In addition, the C-terminal tail is not apparent in crystal structure. Grey shading indicates the part of NS1 ED in complex with the F2/F3 region of CPSF30.

Interestingly, not all NS1 proteins rely equally on the same anti-interferon functions (Kochs et al., 2007a, Geiss et al., 2002, Hayman et al., 2006). Current evidence indicates that NS1 can limit IFN-β induction by both pre-transcriptional (cytoplasmic) and/or co- and post-transcriptional (nuclear) processes. The existence and evolution of two synergistic anti-IFN mechanisms could increase the capacity of some IAVs to adapt to new hosts (Kochs et al., 2007c). In this regard, it is also possible that some virus strains may have lost one or other of these mechanisms, either naturally or during laboratory passage.

1.3.1.2. Pre-transcriptional activities

NS1 counteracts the type I IFN pathway at multiple levels, and notably mediates a pre-transcriptional block of IFN-β induction by interfering with the RIG-I signalling pathway (Mibayashi et al., 2007, Rehwinkel et al., 2010). NS1 is believed to form a complex with RIG-I, although there is no clear consensus on whether this complex involves direct inhibitory contacts between the two proteins or if it is mediated through other interactions (Pichlmair et al., 2006, Guo et al., 2007, Opitz et al., 2007, Mibayashi et al., 2007).
NS1 has also been reported to prevent RIG-I activation indirectly by associating with two positive regulators: the ubiquitin ligases TRIM25 (Gack et al., 2009) and RIPLET (Rajsbaum et al., 2012). The apparent redundancy in NS1's ability to interfere with two activators of

RIG-I could be explained by both viral strain- and host species differences (Rajsbaum et al., 2012). Whether RIPLET and TRIM25 can both ubiquitinate the same residues in the CARD domains of RIG-I remains controversial, however, RIPLET seems to be specifically driving a distinct modification of the N-terminus of the sensor allowing it to change its initial conformational and the subsequent activation by TRIM25 (Oshiumi et al., 2013). According to this model, NS1 could be independently blocking two required regulatory steps for the activation of RIG-I. Furthermore, by binding TRIM25, NS1 can block its oligomerization and subsequent E3 ligase activity on the CARD domains of RIG-I. NS1 residue E96 and E97 are important for this function, and disrupting NS1-TRIM25 interaction via E96A/E97A substitutions have been associated with viral attenuation and higher induction of IFN-β (Gack et al., 2009). Of note, E96/E97 residues are not involved in NS1 inhibition of RIPLET (Rajsbaum et al., 2012).

NS1 has also been reported to block the function of both a constitutively active RIG-I construct lacking its RNA-binding helicase domain, as well as its downstream effector IPS-1 (Mibayashi et al., 2007), however, the mechanism involved is still under investigation.

The dsRNA binding activity of NS1 has also been suggested to participate in RIG-I inhibition by sequestering PAMPs away from the sensor (Talon et al., 2000). The ability to bind different species of RNA was among the first features ascribe to NS1 (Hatada et al., 1999, Hatada and Fukuda 1992). NS1 Dimerization is essential for binding dsRNA (Wang et al., 1999) and the stoichiometry of dimer:dsRNA is 1:1 (Chien et al., 2004). Furthermore, NS1 has previously been reported to interact with the viral polymerase complex (Marion et al., 1997) and with high affinity to dsRNA in the form of vRNA-like panhandle structures (Hatada & Fukuda, 1992; Hatada et al., 1997). Indeed, evidence is accumulating for a number of functional interactions between NS1 and replicating RNPs during infection (Twu et al., 2007). Although one report suggested that NS1 RBD was binding dsRNA with low affinity compared to RIG-I (Chien et al., 2004), several other research groups reported that NS1 had a high affinity for dsRNA,notably in the form of vRNA-like panhandle structures (Hatada & Fukuda, 1992; Hatada et al., 1997). Moreover, mutation of amino-acid R38 and K41 required for NS1 dsRNA binding greatly impairs NS1's ability to block interferon production (Pichlmair et al., 2006). Remarkably, R38/K41 substitutions also abrogate NS1 interaction with RIG-I, TRIM25 and RIPLET (Pichlmair et al., 2006, Gack et al., 2009, Rajsbaum et al., 2012). Thus, it is believed that thanks to its RNA binding properties NS1 can localize into multicomponent and dynamic RNA-protein complexes in which it exerts its inhibition. Furthermore, some pieces of data suggest that such complexes are formed during influenza virus infection (Onomoto et al., 2012).

NS1 has also been shown to prevent activation and nuclear translocation of IRF-3, NF-κB and c-Jun/ATF-2 otherwise essential for IFN-β induction (Ludwig et al., 2002, Talon et al., 2000, Mibayashi et al., 2007, Wang et al., 2000). It was initially proposed that NS1 could inhibit IRF3 activation by binding TRIM-25 and inhibiting RIG-I activation (Gack et al., 2009). However, this hypothesis is debated as some NS1 proteins can bind efficiently TRIM25 (i.e. human seasonal H3N2 viruses) without inhibiting IRF3 activation or IFN-β transcription (Kuo et al., 2010). Of note, the impact of RIPLET-NS1 interaction on IRF3 activation is unknown (Rajsbaum et al., 2012). Work done by Kuo and colleagues also demonstrated that the nature of residue 196 in NS1 ED determines its ability to block IRF3 activation, and subsequently IFN-β induction. These authors demonstrated that viruses harbouring K196 did not block IRF3 activation, whereas viruses harbouring E196 did (Kuo et al., 2010). However, to date the exact mechanisms remain undefined.

While the interaction of IAV NS1 with components of the RIG-I signalling axis are relatively well conserved among virus isolates, the extent and efficacy of inhibition varies significantly among viral strains. Thus, viruses less efficient at controlling the type I IFN pathway at pre-transcriptional level must rely on a different strategy, such as crippling host gene expression at a co- or post-transcriptional level, including that of IFN and IFN-stimulated genes.

1.3.1.3. Co- and post-transcriptional functions

General inhibition of nucleo-cytoplasmic transport of all poly(A)-containing mRNAs was one of the first functions ascribed to NS1 (Fortes et al., 1994, Qiu and Krug, 1994). At the time, it was speculated that global nuclear retention of cellular mRNAs by NS1 was increasing priming of viral mRNA transcription by providing a pool of cap-donors for the viral polymerase complex. However, it is now apparent that blocking cellular mRNA processing goes beyond this and represents an effective mean to limit a number of host-cell processes, including the innate antiviral response.

In eukaryotic cells, the 3'-end of primary transcripts are cleaved upon recognition of a conserved AAUAAA sequence some 10-30 bases upstream of the cleavage site (Figure 1-6). This is the task of the Cleavage and Polyadenylation Specificity Factor (CPSF), a polyprotein complex formed by four subunits (160, 100, 70 and 30 kDa), which recognizes the AAUAAA sequence and binds to nascent mRNAs. It then catalyses an endonucleolytic step and subsequent addition of a poly(A) tail (Wahle and Keller, 1996, Colgan and Manley, 1997).

The smallest components of the CPSF complex, called CPSF30, contains five C3H zinc fingers (Barabino et al., 1997), and in infected cells NS1 bind two of the CPSF30 zinc fingers (F2 and F3). In doing so it inhibits the binding of the whole CPSF complex to pre-mRNA, and subsequently prevents their cleavage and polyadenylation (Nemeroff et al., 1998, Das et al., 2008, Twu et al., 2006). This allows NS1 to inhibit the expression of cellular genes, including those of IFNs and ISGs at a post-transcriptional level. Importantly, viral transcripts are not affected by the repression of CPSF because their polyadenylation is directly catalysed by the viral polymerase (Poon et al., 1999).

NS1 inhibition of CPSF30 may also impact mRNA splicing, in which CPSF is involved (Alonso-Caplen et al., 1992, Alonso-Caplen and Krug, 1991, Fortes et al., 1994, Qian et al., 1994, Li et al., 2001), as well as the nucleo-cytoplasmic transport of mRNAs as it requires polyadenylation (Fortes et al., 1994, Qiu and Krug, 1994). NS1 can also block specifically the nuclear export of fully processed mRNAs that partially escape 3'-end formation inhibition by interacting with the poly(A)-binding protein II (PABPII) through residues 223-237 (Chen et al., 1999) (Figure 1-6).

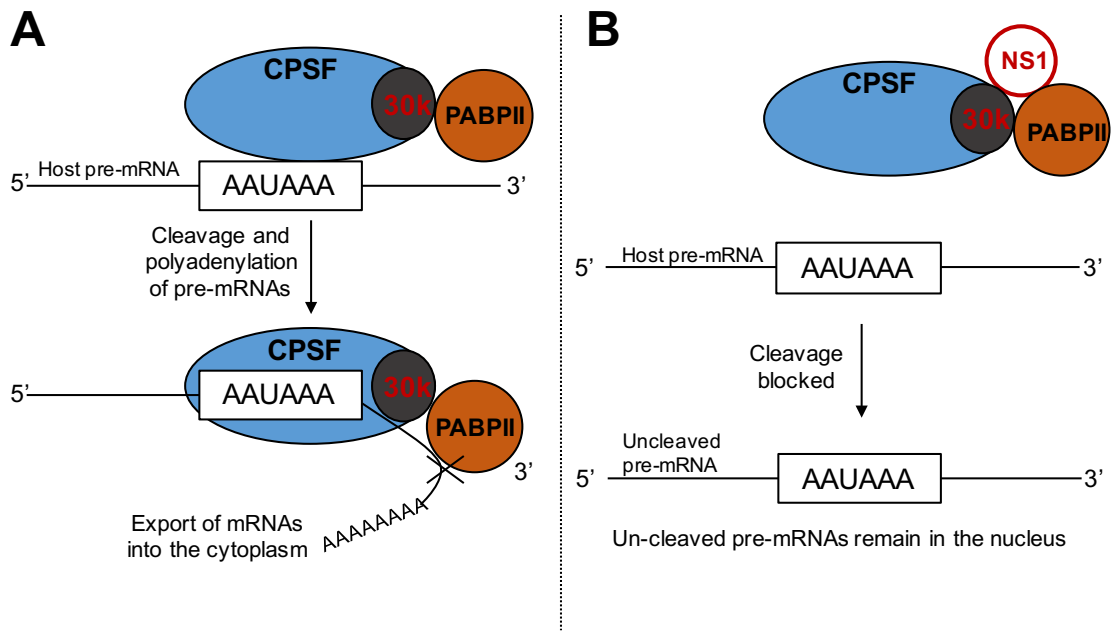


Figure 1-7: NS1 blockade of the cellular mRNA 3'-end processing system.

(A) Normal processing of cellular pre-mRNA by the cleavage and poly-adenylation machinery. **(B)** NS1 prevents the normal processing of cellular pre-mRNA. Mechanism by which IAV NS1 inhibits the 3'-end processing of cellular mRNA in infected cells. Pathway I: the binding of the NS1 protein (and PABII) to the 30 kDa subunit of CPSF blocks the binding of CPSF to the AAUAAA sequence of some cellular pre-mRNA molecules, thereby blocking 3 -cleavage of these pre-mRNAs. The un-cleaved pre-mRNAs remain in the nucleus. Pathway II: CPSF binds to the AAUAAA sequence of other cellular pre-mRNA molecules, despite the binding of the NS1 protein and PABII to the 30kDa CPSF subunit. A short poly(A) sequence is then added to these cleaved pre-mRNAs by PAP in a CPSF-dependent reaction. Subsequent elongation of the short poly(A) sequence is blocked by the binding of the NS1 protein to PABII, resulting in the nuclear accumulation of cleaved pre-mRNAs containing short poly(A) tails.

Residues involved in NS1-CPSF30 interaction have been characterized by mutagenesis (Noah et al., 2003, Twu et al., 2007, Li et al., 2001, Kochs et al., 2007b) and by solving the three-dimensional structure of NS1 ED in complex with the CPSF30 F2F3 region (Das et al., 2008). A patch of highly conserved hydrophobic residues (184 to 188) was shown to embed into a pocket on the F2F3 loop of CPSF30, with G184 standing in a central position. Interestingly, G184A substitution abrogates NS1 binding to CPSF30 (Das et al., 2008). In addition, based on crystal structures NS1 ED dimerization is incompatible with CPSF30 binding (Kerry et al., 2011b), and W187 plays a pivotal role in driving both interactions (Kerry et al., 2011a, Aramini et al., 2011). Such incompatibilities are likely to play a role on the spatial-temporal regulation of NS1 functions.

Additional residues are important for stabilizing NS1-CPSF30 interaction, and notably F103, M106, K108, D125, L144, and D189, although they are not strictly part of the binding interface (Kochs et al., 2007a, Li et al., 2001, Noah et al., 2003, Twu et al., 2007, Twu et al., 2006, Das et al., 2008, Hale et al., 2010). Glu-96 may also be functionally significant (Shimizu et al., 1999).

Of note, the polymerase subunits (PB1, PB2, PA) and the NP protein are important to stabilise NS1-CPSF interactions in a strain-specific manner (Twu et al., 2006, Kuo et al., 2010).

Although CPSF30 and its F2F3 loop are relatively well conserved among mammals, as seen in the alignment of human and equine CPSF30 (Figure 1-7), variations in CPSF30 binding seem to arise commonly during viral adaptation to different hosts (Hossain et al., 2008, Brown et al., 2001).

Figure 1-8: Human and equine CPSF30.

Alignment of human (top) and equine (bottom) CPSF30 amino acid sequences (ENSECAG00000007525 and ENSG00000160917, respectively). Red residues highlight amino acid differences. Red square indicates the F2F3 loop of CPSF30 bound by some IAV NS1 proteins.

In general, mouse-, egg- and swine adapted viruses have NS1s that bind weakly to CPSF30, whereas those from human-circulating strains are strong binders, with the exception of the human pandemic H1N1 virus from 2009 (2009pH1N1), whose NS1 is of swine origin. In the case of avian strains both NS1 phenotypes can be found (Twu et al., 2006, Hale

et al., 2010). For PR8/NS1, the lack of CPSF30 binding is due to substitutions of residues 103 and 106 (Kochs et al., 2007b), while for 2009pH1N1/NS1 the defect in CPSF30 binding is due to K108R, D125E, D189G substitutions (Hale et al., 2010). Interestingly, restoring NS1 binding to CPSF30 slightly decreased replication and virulence of the 2009pH1N1 virus in mice.

While inhibition of CPSF30 is presumed to non-specifically block general gene expression, some NS1 proteins may have evolved mechanisms to repress the expression of specific gene sets. This is the case of human seasonal H3N2 viruses isolated after 1989. As previously mentioned, the NS1 protein of these viruses reverted back from 237 to 230 amino acids (Hale et al., 2008c) As a result, their C-terminal residues became 226-ARSKV-230, a very similar motif to the 'ARTKQ' motif found on cellular histone H3 (H3). Consequently, the tail of these NS1 proteins can act as histone-mimics. By doing so, these NS1 proteins could repress expression of specific cellular genes by competing with H3 for binding to the Polymerase Associated Factor 1 (PAF1)-transcription-elongation complex (Marazzi et al., 2012). Interestingly, PAF1-regulated genes include those involved in the IFN response.

Finally, the NS1 protein of the A/WSN/1933-H1N1 (WSN) strain has been reported to interact with several components of the cellular mRNA nuclear export machinery (NXF1, p15, Rae1, E1B-AP5, Nup98) and block their function (Satterly et al., 2007). As different cellular mRNAs use different export factors to translocate to the cytoplasm (Satterly et al., 2007), inhibition of cellular mRNA export may give some specificity for inhibition of cellular gene expression to WSN/NS1.

1.3.1.4. Inhibition of antiviral gene products

NS1 IFN antagonism goes beyond the pre- and post-transcriptional repression of IFN genes. Indeed, NS1 also blocks directly the antiviral effects of some ISG-products, such as PKR (Min et al., 2007) and the OAS/RNAse L-pathway (Min and Krug, 2006). As mentioned previously (see section 1.2.4), both OAS and PKR are key regulators of viral transcription/translation processes, and play additional roles in other innate defences, such as IFN-β induction and the host apoptotic response to infection (Silverman, 2007). Importantly, thanks to this downstream activity of NS1, viral replication could take place in cells that have been set in an active antiviral state by a prior IFN treatment (Newby et al., 2007, Min and Krug, 2006).

IAV was found to inhibit PKR activity in early studies (Katze et al., 1986, Katze et al., 1988), and it was postulated to antagonize it via two mechanisms: the regulation of p58IPK function, a regulator of PKR activity (Melville et al., 1999, Goodman et al., 2007), and by the action of NS1. Several studies have suggested that NS1 exerted its PKR inhibitory action by sequestering dsRNA away from PKR (Hatada and Fukuda, 1992). However, the affinity of NS1 for dsRNA is much lower than that of PKR, thus out-competition seems unlikely (Li et al., 2006). Furthermore, a R38A-RNA-binding defective NS1 protein was shown to limit PKR activation efficiently in response to dsRNA activation (Li et al., 2006). Based on domain mapping studies, it has been proposed that NS1 binds to a linker region in PKR, which requires residue 123 to 127 (Li et al., 2006, Min et al., 2007), and thereby prevents a conformational change that is normally required for release of PKR auto-inhibition (Li et al., 2006).

The proposed mechanism of NS1 inhibition of the OAS-RNAse L pathway is the sequestration of dsRNA away from OAS, as in this case the dsRNA binding affinity of OAS is low enough to be outcompeted by NS1 (Min and Krug, 2006). Furthermore, no direct interaction has been described between NS1 and OAS and/or RNAse L. Given the role of RNAse L in augmenting the production of IFN-β (Silverman, 2007), it is possible that NS1-mediated OAS inactivation also contributes to suppression of IFN-β synthesis (Donelan et al., 2004, Talon et al., 2000).

1.3.1.5. Activation of PI3K pathway and regulation of the apoptotic response

Phosphoinositide-3-kinases (PI3K) are ubiquitously expressed and highly conserved cytoplasmic heterodimeric enzymes (Engelman et al., 2006). Cellular responses involving PI3K signalling include survival, proliferation, trafficking, and regulation of the immune function (Engelman et al., 2006). Upon activation, PI3K catalyses the formation of PIP3 (phosphatidylinositol (3,4,5) triphosphate), a membrane-embedded second messenger, which acts as docking platform for proteins involved in signalling, the best characterized being the Serine/Threonine kinase Akt.

Influenza A virus is known to activate PI3K signalling twice during infection, as detected by the phosphorylation state of Akt (Ayllon et al., 2012, Ehrhardt and Ludwig, 2009). A first early and transient PI3K activation peak takes place during viral attachment and entry, probably related to virus uptake by endocytosis (Ehrhardt et al., 2006), whereas a second and sustained activation appears 2-3h post-infection and is due to NS1 (Hale et al., 2006,

Shin et al., 2007). The heterodimeric PI3K is kept at its basal state due to inhibitory contacts exerted by the regulatory subunit (termed p85) over the catalytic one (termed p110). This inhibition is mediated by two SH2 domains on p85. NS1 binds to p85 and displaces the N-terminal SH2 domain from its position on the heterodimer, thus releasing the inhibition over p110 (Hale et al., 2008b). A highly conserved tyrosine residue at position 89 in NS1 settles in the centre of the NS1-p85 interphase (Shin et al., 2007), and substitution of this amino acid for a phenylalanine (Y89F) abrogates p85 binding and PI3K/Akt activation (Hale et al., 2006).

1.3.1.6. Regulation of viral RNA and protein synthesis

Upon infection, IAV takes over the cellular machinery by preferentially translating its own viral products over the host's. NS1 contributes to this process by shutting off cellular gene expression, as described above, and by directly enhancing the translation of viral mRNAs. The first 113 amino acids of NS1 are important for this function (Marion et al., 1997), and so is its interaction with proteins involved in eukaryotic translation, such as the elongation initiation factor 4GI (eIF4GI) (Aragon et al., 2000), polyadenylate binding protein I (PABPI) (Burgui et al., 2003) and hStaufen (Falcon et al., 1999). It has been proposed that NS1 selectively relocalises these factors to the 5'-UTRs of the viral genomic segments in cytoplasmic polysomes (Krug and Etkind, 1973).

NS1 can also affect vRNA temporal regulation. Viral genes have traditionally been divided into early and late stage genes, with the NS and NP genes being selectively expressed early and all the eight genomic segments showing higher expression levels at latter stages (Shapiro et al., 1987, Skehel, 1973). Alanine substitutions of NS1 amino acids 123 and 124 resulted in a temporal deregulation of vRNA synthesis, with late genes being generated earlier and in greater numbers (Min et al., 2007). Further evidence that NS1 plays a role in regulating viral RNA synthesis came from recombinant PR8 viruses carrying deletions of the whole NS1 ED, which specifically impair the mRNA levels of HA (Maamary et al., 2012). Although the mechanism involved is unknown, it has been suggested to be related to NS1 interaction with the vRNP through its direct binding to NP (Robb et al., 2011). Relatively recent data also suggested that NS1 inhibition of DDX21 might be key for this process. Indeed, DDX21 is a cellular helicase able to block the viral polymerase complex formation and can also interact with both the NS1/CPSF30 complex and PB1 (Chen et al., 2014).

1.3.1.7. Interaction with other host factors

In full length NS1 proteins (230 amino acids), a PSD-95 Discs-large ZO-1 (PDZ)-binding motif (PBM) has been identified from residue 227 to 230 (Obenauer et al., 2006). PBMs are important protein recognition signals relevant in the assembly of multi-component signalling complexes (Javier and Rice, 2011). They recognize and bind short C-terminal peptide motifs of 4-5 amino acids, called the PDZ domain ligand (PL). The most common PBM found in avian IAV NS1 proteins is "ESEV/EPEV", while in human IAV NS1s the most common motif found is "RSKV/RSEV". Although scarce information is available for NS1 of IAVs isolated from other species, sequence analysis revealed that full length NS1 EIVs possessed an avian PBM (Barba and Daly, 2016). The avian PBM has been shown to interact with up to thirty human PDZ domain-containing proteins, whereas human NS1 proteins could not (Obenauer et al., 2006). By interacting with PDZ domain-containing proteins, IAVs may disrupt certain cellular pathways and cause virulence. Avian motifs have been found to interact with PDZ domain-containing proteins, such as Scribble and Dlg1 involved in cellular junctions and apoptotic responses (Golebiewski et al., 2011, Javier and Rice, 2011, Liu et al., 2010). The relevance of this avian-specific motif is unclear, and has been shown to increase replication of some IAV strains, e.g. WSN (Jackson et al., 2008) and avian H7N1 viruses (Soubies et al., 2010), while not affecting others, e.g. A/VN/1203/04 H5N1 HPAI virus (Zielecki et al., 2010). The specific targets, mechanisms and strain specificity of this avian PBM remain to be fully identified.

1.3.1.8. Intracellular distribution

NS1 was originally described as a nuclear protein and is predominantly found in the nucleus of infected cells, especially at early times post-infection (Lazarowitz et al., 1971). The nucleo-cytoplasmic location of NS1 can be regulated by a monopartite nuclear location signal (NLS) (amino acids 35-41), which overlaps with critical residues for dsRNA binding (R38 and K41) and is mediated through interaction with importin-α (Greenspan et al., 1988). NS1 also possesses a nuclear export signal (NES) found in its ED, within amino acids 138 to 147 (Li et al., 1998). In addition, a bipartite NLS can also be found in 237 residue-long NS1 proteins (position 219 to 232) (Melen et al., 2007). As some NS1 activities must take place in the nucleus (e.g. interference with CPSF30 and PABPII function), while others should happen in the cytoplasm (e.g. inhibition of RIG-I and activation of PI3K), the subcellular location of the protein must play a role in determining which functions are executed during infection.

1.3.1.9. Post-translational modifications

Post-translational modification is a common mechanism to regulate multi-functional proteins. As such, several studies have addressed which of these modifications can affect NS1. To date, NS1 has been shown to be modified by phosphorylation and by coupling to the ubiquitin-like proteins ISG15 and SUMO.

Early studies showed that the WSN/NS1 was phosphorylated soon after infection, probably in the nucleus (Privalsky and Penhoet, 1978, Privalsky and Penhoet, 1981). Structural analysis suggested different serine and threonine residues were available for phosphorylation (Bornholdt and Prasad, 2006), and residues T215 (Hale et al., 2009), S42 and S48 (Hsiang et al., 2012) were later identified. Only phosphorylation of serine 42 by cellular PKCα, has been shown to play a relevant role for the virus, as its substitution on Ud/NS1 affected dsRNA binding and consequently replication (Hsiang et al., 2012). Serine 48 is not completely conserved and its substitution has no impact on viral replication (Hsiang et al., 2012); while Threonine 215, which can be phosphorylated in vitro by CDK and ERK kinases, is a distinctive feature of human-adapted strains, and avian isolates usually display a proline at this position (Hale et al., 2009).

Several lysine residues can be ISGylated in NS1, and K41 seems to be the main target during infection (Tang et al., 2010). As mentioned earlier, K41 is involved in both dsRNA binding and nuclear import. Different studies have described covalent binding of NS1 to ISG15 (Zhao et al., 2013, Tang et al., 2010, Kuo et al., 2010). ISG15 modification of K41 impairs NS1 interaction with the karyopherin without affecting its dsRNA-binding properties. The proposed mechanisms for ISG15-based suppression of the NS1 function, are prevention of PKR binding, abrogation of RBD dimerization and overall subpar antagonism of the IFN system (Tang et al., 2010). As with most NS1-associated features, ISGylation has a strong strain-specific component, and naturally occurring K41R substitutions have been described in certain human H3N2 virus strains without affecting their anti-IFN function (Tang et al., 2010, Zhao et al., 2013).

As mentioned above, another ubiquitin-like protein that can be conjugated to NS1 is SUMO (Pal et al., 2011, Xu et al., 2011, Santos et al., 2013). These proteins can regulate function, stability, location, and interactivity of numerous proteins, both of cellular and viral origins (Everett et al., 2013). NS1 can be modified by the three main isotypes of SUMO, such as SUMO 1 and SUMO 2/3 (Santos et al., 2013). Unsurprisingly, NS1 SUMOylation is

strain-dependent, and NS1 SUMOylation is even required for optimal IFN antagonism for some IAV strain (e.g. PR8) (Xu et al., 2011).

1.3.1.10. NS1 contribution to virulence

Viruses expressing NS1 proteins with large C-terminal deletions or a total lack of NS1 altogether are generally attenuated in mammalian and avian cells and are strong inducers of type I IFN (Solorzano et al., 2005, Cauthen et al., 2007).

The outcome of disease also depends on the balance between pro-inflammatory cytokine production and the ability of NS1 to overcome it (Hyland et al., 2006). The H5N1 highly pathogenic avian influenza (HPAI) virus (HK97) responsible for an outbreak in 1997 in human populations was associated with hyper-cytokinemia, which was consistent with the detailed post-mortem results of individuals that died during the 1997 H5N1 HPAI outbreak. This virus was also resistant to the antiviral effects of type I IFN and TNF-α (Cheung et al., 2002, Seo et al., 2002). It was later shown that this cytokine imbalance required E92 in NS1 (Seo et al., 2002). However, E92 seemed to be specific to these H5N1 HPAI viruses from 1997, and it has never been found in NS1 of other IAV subtypes. This highlights the species-specificity of NS1 and its impact on virulence.

NS1 species-specific activities and correlation with virulence was also illustrated by Basler and colleagues' work, who exchanged the NS segment of the mouse-lethal WSN strain for that of the 1918pH1N1. The recombinant virus obtained replicated well in tissue-culture but was attenuated in mice compared to wild-type WSN (Basler et al., 2001). It was speculated that attenuation of the recombinant virus may be due to the human origin of the 1918pH1N1/NS1, which is adapted to function well in human cells, but is unable to work optimally in murine cells.

Another mechanism by which NS1 may affect virulence is by binding to and interfering with cellular signalling proteins via its PBM (Obenauer et al., 2006). Indeed, the introduction of an avian PBM into WSN/NS1 increased WSN virulence in mice, and infections were characterized by a severe loss of body weight, decreased survival, severe alveolitis and increased viral spread in the lung. This work supported the hypothesis of Obenauer and Colleagues indicating that avian NS1 proteins, when present in human cells, may interact

with PDZ domain-containing proteins, disrupt certain cellular pathways, and increase virulence (Obenauer et al., 2006).

The contribution of host adaptation and NS1 species-specific functions to viral pathogenesis has been of interest to the scientific community for many years. However, little is known about the contribution of NS1 evolution to EIV pathogenesis.

1.3.2. Other viral players

NS1 is not the only viral protein capable of restraining the innate immune system. Indeed, IAVs encode several other proteins that counteract the interferon pathway either directly or indirectly. For example, both PB2 and PB1-F2 limit the production of IFN-β through association with IPS-1 (Iwai et al., 2010, Varga et al., 2012), especially variants containing D9 in PB2 and N66 in PB1-F2 (Varga et al., 2011) (Long and Fodor, 2016, Varga et al., 2012). Some virus strains possess highly efficient polymerase complexes that allow the virus to outrun the interferon effects and contribute to immune evasion of IAVs (Grimm et al., 2007). The viral polymerase is also involved in cap-snatching of host mRNAs and thereby reduces host cell gene expression, including that of IFN (Nakhaei et al., 2009). The recently discovered PA-X viral protein is also able to repress cellular gene expression, and notably genes involved in regulating the initiation of the cellular immune response (Feng et al., 2016). The NP protein has also been shown to interfere with PKR function by promoting its inhibition by p58IPK (Sharma et al., 2011). Furthermore, by encapsidating the vRNA, NP is likely to reduce the formation of dsRNA, which could otherwise activate viral sensors and lead to activation of the anti-viral response. In addition, since most PRRs are located in the cytoplasm, the nuclear replication strategy of IAV also prevents recognition of vRNA by cytosolic PRRs. Also, IAV can directly infect and kill NK cells (van de Sandt et al., 2012), or reduce its recognition of virus-infected cells due to the gradual mutation of glycosylation sites of the HA protein (Medina et al., 2013).

Besides interfering with innate signalling, IAVs are also able to counteract cells of the adaptive immune system. For example, infection of monocytes impairs their ability to differentiate into mature dendritic cells, which indirectly limits induction of virus-specific CD8+ T cell responses (van de Sandt et al., 2012). In addition, antigenic drift or shift allow IAV to evade antiviral humoral responses and may affect antigen presentation in the context of an MHC molecule, hence preventing T cell activation (Bryant et al., 2009, van de Sandt et al., 2012).

Cross-species jumps from avian reservoirs to mammalian hosts are usually associated with sporadic outbreaks, but on rare occasions emergent viruses manage to establish a lineage in the new host species. The immune pressure exerted by the new host on the emergent virus forces it to evolve and adopt new strategies to evade immunity in order to survive in nature. Understanding the biological mechanisms that allow successful inter-species transmission and adaptation to mammals is crucial to develop the theoretical tools required to predict and/or control emergence of new viruses in humans and animals. H3N8 EIV represents an interesting model to study the dynamic of within-host variation of avian-origin IAVs, as this virus has emerged from a bird reservoir in 1963 and has now circulated for more than fifty years in horse populations despite the availability of vaccines with barely any reassortment events. Importantly, forty years post emergence this virus has jumped the species barrier and established a new lineage in dogs, which suggest that H3N8 EIV has successfully adapted to mammals. Evidence of evolution of EIV virulence factor NS1 also exists, however, little is known about its impact on EIV infection phenotype and adaptation to horses. This virus represents then a unique tool to study NS1 contribution to virus-host interaction and successful adaptation to mammals.

Chapter 2

Aims

2. Aims

The overall aim of this project was to characterise the molecular evolution of NS1 during the post-transfer adaptation process of H3N8 EIV, utilizing reverse genetics and site-directed mutagenesis techniques, in association with a phylogenetic analysis and high-throughput sequencing technologies.

Specifically, I aimed to:

(i) To identify molecular determinants of EIV NS1 evolution.

(ii) To analyze the role of specific mutations on NS1 function at the protein level.

(iii) To confirm the role of the NS1 gene in EIV infection phenotype.

(iv) To evaluate the impact of NS1 evolutionary markers on EIV infectivity in mammalian cells.

(v) To identify which advantages are associated with NS1 evolutionary markers E186K and C-terminal truncation.

(vi) To compare the early genes expression pattern of equine cells infected with two evolutionary distinct EIVs.

(vii) To evaluate the impact of NS1 evolutionary markers on virus-host interaction.

Chapter 3

Materials and Methods

3. Materials and Methods

3.1. Generation of mammalian expression constructs for NS and untagged NS1

3.1.1. Virion RNA extraction

vRNA was extracted from 140μl of the corresponding viral stocks using the QIAamp® Viral RNA Mini kit (Qiagen) according to the manufacturer protocol. AVL buffer and absolute ethanol were used to inactivate the virus. Samples were then loaded into RNeasy mini-spin columns and RNA was bound to the column membrane by centrifugation at 8000xg for 30 seconds. The membrane was washed once with 700μl of Buffer WT1 and twice with 500μl Buffer RPE (provided with the kit). The RNA was then eluted into a new microcentrifuge tube with 50μl RNase-free water pipetted directly onto the silica-gel membrane of the column and centrifuged at 13000xg. RNA eluted from the columns was quantified using a spectrophotometer (NanoDropTM 1000, Thermo Scientific). The ratio of absorbance at 260nm and 280nm, and 260nm and 230nm were used to assess the purity of RNA. A ratio of 2.0 for both was considered optimal for RNA. Samples were then stored in RNAse-free, low binding tubes at -80ºC until further use.

3.1.2. Retro-transcription

Reverse transcription was performed using Uni12 primer (5'-GCCGGAGCTCTGCAGATATCAGCRAAAGCAGG-3') (Hoffmann et al., 2001), 500ng of total RNA, and SuperScript III Reverse Transcriptase (Invitrogen) following the manufacturer's protocol, and the cycling parameters provided below (Table 3-1).

Table 3-1: Retro-transcription cycling parameters

Steps	Temperature	Duration
Annealing	65ºC	5 minutes
Incubation	ice	1 minute
First strand cDNA synthesis **(1 cycle)**	55ºC	30 minutes
Inactivation	70ºC	15 minutes

3.1.3. PCR

Amplification of the NS1 gene or the full NS segment was performed using specific primers (Table 3-2 and 3-3, respectively) by Polymerase Chain Reaction (PCR) with 100ng of cDNA, and the Pfu-Ultra II fusion HS DNA polymerase (Agilent) following the manufacturer's protocol. Primers designed to amplify the full NS segment or the NS1 gene possessed an EcoRI restriction site at the 5'-end of the forward primer and an XhoI restriction site at the 5'-end of the reverse primer. The amplification reactions were performed in the Applied Biosystem® 2720 thermocycler and indicated cycling parameters (Table 3-4). Amplified products were examined by electrophoresis on 1% agarose gel.

Table 3-2: Primers for NS1 gene cloning

Viruses	Primers
U/63, M/63, F/79, S/89, K/91, LP/95, K/95, K/99, K/02, N/03, O/03, M/13	FW: 5'-AGGCGAATTCGCCACCATGGATTCCAACACTGTGTCAAGCTTTCAG-3'
SP/69	FW:5'-AGGCGAATTCGCCACCATGGATCCCAACACTGTGTCAAGCTTTCAG-3'
U/63, M/63, SP/69, F/79	RV: 5'-CGGGCTCGAGTCAAACTTCTGACTCAATTGTTCTCG-3'
S/89, LP/95, K/91	RV: 5'-CGGGCTCGAGTCAAACTTCTGGCTCAATTGTTCTC-3'
K/99, K/02, O/03, N/03, M/13	RV: 5'-CGGGCTCGAGTCATTTCTGCTTTGAAGGGAAT-3'
K/95	RV: 5'-CCCGCTCGAGTCAAATTTCTGGCTCAATTGTT-3'

Table 3-3: Primers for NS segment cloning

Viruses	Primers
U/63, O/03	FW: 5'-TATTGAATTCCGTCTCAGGGAGCAAAAGCAGGGTG-3'
U/63, O/03	RV: 5'-CGGTCTCGAGCGTCTCGTATTAGTAGAAACAAGGGTGTTTTTTT-3'

Table 3-4: Cycling parameters for NS and NS1 gene cloning

Steps		Temperature	Duration
Initial activation		95ºC	1 minute
3 steps cycling (30 cycles)	Denaturation Annealing Extension	95ºC 65ºC 72ºC	20 seconds 20 seconds 30 seconds
Final extension		72ºC	3 minutes

3.1.4. Restriction digest

NS1 and NS PCR products were purified with the QIAquick PCR Purification Kit (Qiagen), following the manufacturer protocol. To clone NS1 genes in pcDNA3.1+ or pCAGGS plasmid and NS segments in pcDNA3.1+, NS1 or NS PCR products and pCAGGS or pcDNA3.1+ plasmids were digested with XhoI and EcoRI (Promega) for 3 hours at 37ºC, following the manufacturer protocol. To clone NS segments in pDP2002, NS PCR products and pDP2002 plasmid were digested with BsmBI restriction enzyme (Promega) for 3 hours at 37ºC, following the manufacturer protocol. Digested plasmids and PCR products were then examined on 1% agarose gel electrophoresis, and the bands of interest were extracted and purified with QIAquick Gel Extraction Kit (Qiagen). DNA purity and concentrations were analysed by Nanodrop.

3.1.5. DNA ligation

The T4 DNA Ligase (Promega) was used to ligate NS PCR products into pcDNA3.1+ or pDP2002, and NS1 PCR products into pCAGGS with a vector:insert ratio of 1:5. The ligation was carried out at 4ºC overnight.

3.1.6. Bacterial transformation and plasmid preparation

One Shot TOP10 Chemically Competent E. coli (Invitrogen) were transformed following the manufacturers protocol with 5μl of ligation product. Transformed colonies were grown in 0.1mg/ml Ampicillin-containing Luria-Bertani (LB) agar plates (Fast-Media® Base Agar, Invivogen; ampicillin, Sigma-Aldrich), and incubated at 37ºC overnight. Ten colonies per construct were selected and cultured in 5ml of LB-Broth supplemented with 0.1mg/ml Ampicillin for 24h at 30°C in a shaker incubator at 220rpm.
Transformation of XL-10 Gold Ultracompetent Cells (Agilent) was performed using 2μl of ligation product. The bacteria were then cultured on 0.1mg/ml Ampicillin-containing LB agar plates overnight at 37ºC. Bacterial cultures containing the recombinant plasmid were purified using the QIAprep Spin Miniprep Kit (Qiagen) according to the manufacturer's protocol. Purified plasmids were quantified using NanoDrop Spectrophotometer.

3.1.7. Agarose gel electrophoresis

For agarose gel preparation, 10X TAE buffer (0.4M Tris, 1.142% (v/v) acetic acid, 0.01M EDTA, pH adjusted to 8, Sigma-Aldrich) was generated and diluted to 1X in H_2O prior to use. 0.8-1% (w/v) molecular grade agarose (Bioline) diluted in 1X TAE. 1Kb DNA ladder (Promega). Gels were stained with ethidium bromide (10mg/ml, Biotechnology grade, AMRESCO).

3.1.8. Quick-change PCR

Silent mutations were introduced by site-directed mutagenesis into the splice acceptor site (intron/exon junction: 525-CCAC↓GA-530 changed to 525-CCCGGG-530) of the NS1 pcDNA3.1+ inserts generating NS1 splice-acceptor mutant (SAM) to avoid the expression of NEP. To engineer NS1 and NS mutants for residue 186, Lysine (K)-to-Glutamic acid (E) (codon 186 'AAA' changed for 'GAA') or E-to-K substitutions (codon 186 'GAA' changed for 'AAA") for O/03 and K/95 or U/63, respectively, were introduced into the NS1 SAM- or NS-pcDNA3.1+ inserts using the PFU-Turbo DNA polymerase (Agilent) and specific sets of primers (Table 3.5), and cycling parameters (Table 3.6). To engineer NS1 C-terminus mutant in NS1 or NS inserts, Stop Codon-to-Arginine (R) (codon 220 'TGA' replaced by 'CGA') or Arginin-to-Stop Codon (codon 220 'CGA' to 'TGA') for O/03 and K/95 or U/63, respectively, were introduced into the NS1 SAM- or NS-pcDNA3.1+ as described above. A similar strategy was used to engineer U/63 and O/03 NS1 112 mutants (Alanine (A)-to-Threonine (T) codon 'GCA' changed for 'ACA" or T-to-A codon 'ACA' changed for 'GCA').
Purified recombinant plasmids carrying either NS1 SAM, NS1 or NS 186 and/or C-terminus mutants were used as templates to amplify the corresponding inserts, which were then sub-cloned into either pCAGGS for NS1 or pDP2002 for NS. This was performed using the PFU-Ultra specific primers and cycling parameters (Table 3.5 and 3.6, respectively).

Table 3-5: Primers for site-directed mutagenesis in NS or NS1 genes

Viruses	Mutation	Primers
U/63, M/63, SP/69, F/79, S/89, K/91	SAM	FW: 5'-CATTGCCTTCTCTTCCCGGGCATACTAATGAGGATGTC-3' RV: 5'-GACATCCTCATTAGTATGCCCGGGAAGAGAAGGCAATG-3'
LP/95, K/95, K/99, K/02, N/03, O/03, M/13	SAM	FW: 5'-CATTACCTTCTCTTCCCGGGCATACTAATGAGGATGTC-3' RV: 5'-GACATCCTCATTAGTATGCCCGGGAAGAGAAGGTAATG-3'
U/63	A112T	FW: 5'-CATGCCCAAGCAAAAAGTAACAGGCTCCCTATGTATAAG-3' RV: 5'-CTTATACATAGGGAGCCTGTTACTTTTTGCTTGGGCATG-3'
O/03	T112A	FW: 5'-CATGCCCAAGCAAAAAGTAGCAGGCTCCCTATGTATAAG-3' RV: 5'-CTTATACATAGGGAGCCTGCTACTTTTTGCTTGGGCATG-3'
U/63	E186K	FW: 5'-CCTCATCGGAGGACTTAAATGGAATGATAACAC-3' RV: 5'-GTGTTATCATTCCATTTAAGTCCTCCGATGAGG-3'
O/03 K/95	K186E	FW: 5'-CCTCATCGGAGGACTTGAATGGAATGATAATAC-3' RV: 5'-GTATTATCATTCCATTCAAGTCCTCCGATGAGG-3'
U/63 K/95	R220*	FW: 5'-CCAAAGCAGAAATGAAAAATGGCGAGAACAATTGAG-3' RV: 5'-CTCAATTGTTCTCGCCATTTTTCATTTCTGCTTTGG-3'
O/03	*220R	FW: 5'-CCCTTCAAAGCAGAAACGAAAAATGGAGAGAAC-3' RV: 5'-GTTCTCTCCATTTTTCGTTTCTGCTTTGAAGGG-3'

Table 3-6: Cycling parameters for site-directed mutagenesis in NS or NS1 genes

Steps		Temperature	Duration
Initial activation		95°C	2 minutes
3 steps cycling (25 cycles)	Denaturation Annealing Extension	95°C 60°C 72°C	30 seconds 1 minute 11 minutes
Final extension		72°C	10 minutes

3.1.9. Plasmids sequencing

The sequence of each pCAGGS- or pcDNA3-NS1 inserts and pDP2002- and pcDNA3.1-NS inserts were confirmed by Sanger sequencing and compared to those available in the NCBI Influenza Database (Appendices 3 to 6 and 32 to 33)

3.2. Luciferase assays

For analysis of IFN-β and ISRE promoter activation and control of general gene expression, HEK293T cells (12-well plate format, $2.5x10^5$ cells/well) were transiently co-transfected using TransIT-LT1 transfection reagent (Cambridge Biosciences), with either

50ng of pIFN-β-FF-Luc or 50ng of pISRE-FF-Luc (reporter plasmids encoding Firefly luciferase [FF-Luc] under the control of the IFN-β promoter or the ISRE promoter, respectively), 50ng of a plasmid constitutively expressing Renilla luciferase (pREN-Luc) under the SV40 promoter (kindly provided by Benjamin G. Hale), as well as 3 different amounts (1000ng, 200ng or 20ng) of the indicated NS1 SAM-pCAGGS expressing plasmids (or empty pCAGGS,1000ng). The total amount of transfected plasmid DNA was kept constant (1000ng) with empty plasmid (pCAGGS). After 24h of transfection, cells were infected with 50 hemagglutinating units (HAU) of SeV for 18h, then lysed with 250µl of passive lysis buffer (Promega). IFN-β-FF-Luc, ISRE-FF-Luc and REN-Luc activities were measured using the Dual-luciferase reporter assay system (Promega), as directed by the manufacturer protocol and a Glomax® 20/20 luminometer (Promega). All transfections were carried out in triplicates, and experiments were repeated independently three times.

3.3. Co-immunoprecipitation of NS1 with CPSF30

The equine CPSF30 was cloned as previously described for the human CPSF30 (Kochs et al., 2007). The equine CPSF30 gene was amplified by RT-PCR from equine cells using oligo d(T) and PCR with specific primers (Forward 5'-ATGCAGGAAATCATCGCCAG-3', reverse primer: 5'-CCTTTCTCAGTGGACAGTGA-3'). The RT-PCR product was then cloned into the pCAGGS HA-COOH plasmid as described previously [35]. NS1 proteins were synthesized in vitro using pcDNA3.1+ plasmid and the TNT7 transcription/translation kit (Promega) following the manufacturer's recommendations. HEK293T cells ($1.5x10^6$ cells/well, 6-well format, triplicates) were transiently transfected with 2µg/well of an HA-equine CPSF30-pCAGGS-expressing plasmid. At 48 hours post-transfection (hpt), cells were lysed in 20mM Tris-HCl (pH 7.5), 100mM NaCl, 0.5mM EDTA, 5% glycerol, 1% Triton X-100 supplemented with a complete mini protease inhibitor cocktail (Pierce). Cleared cell lysates were incubated overnight at 4°C with the *in vitro*-synthesized O/03 and revertant NS1 proteins and 20µl of an anti-HA affinity resin (Sigma-Aldrich). After extensive washing, precipitated proteins were dissociated from the resin using Laemmli buffer + β-mercaptoethanol. Total proteins from cells lysates were then analysed by SDS-PAGE and Western blotting (as described below) using a primary specific rabbit polyclonal antibody against NS1 (Genscript) or HA-tag for CPSF30 (Sigma-Aldrich) (Table 3-8).

3.4. Cell culture

3.4.1. Cell maintenance

Sterile standard tissue culture conditions were maintained in all experiments. Madin-Darby Canine Kidney (MDCK, ATCC CCL-34) and Human Embryonic Kidney cells (HEK293T, ATCC CRL-11268) were grown at 37ºC, 5% CO_2 in Dulbecco's modified Eagle's medium (DMEM) high glucose, GlutaMax and pyruvate supplemented (ThermoFisher Scientific), 10% of Fetal Calf Serum (Gibco Life Technologies), and 1% PS (penicillin, 100 units/ml; streptomycin, 100μg/ml; Gibco Life Technologies).
Equine Dermal fibroblasts (E.Derm, NBL-6; ATCC CCL-57) were grown at 37ºC, 5% CO_2 in DMEM high glucose, GlutaMax and pyruvate supplemented, 15% of Fetal Calf Serum, 1% Non-Essential Amino acids (Gibco, Life Technologies), 1% PS.

3.4.2. Subculture

Cells were routinely grown in 150cm2 flasks and passaged when reaching confluence. For propagation, cell monolayer was rinsed with phosphate buffered saline without calcium (PBS, 137mM NaCl, 15mM KCL, 10mM Na2HPO4/KH2PO4, pH 7.4, ThermoFisher Scientic), and incubated with 5ml of 0.05% trypsin solution containing 200mg/L Ethylenediaminetetraacetic acid (EDTA) (ThermoFisher Scientific) for 3 minutes at room temperature for E.Derm and HEK293T cells and 20 minutes at 37ºC, 5% CO_2 humidified atmosphere for MDCK cells. Once detached cells were re-suspended in growth medium and pelleted at 200xg for 5 minutes. The medium was then discarded and replaced by fresh growth medium, and cells were seeded in new sterile 150cm2 flasks or used for assays.

3.4.3. Long-term storage

For long term storage in liquid nitrogen, cells were harvested and pelleted at 200xg for 5 minutes and re-suspended in 90% (v/v) Fetal Calf Serum, 10% (v/v) dimethyl sulfoxide (DMSO) (ThermoFisher Scientific). The suspension was divided in 1ml aliquots ($1x10^6$ cells/vial) in 2ml cryogenic tubes (ABC Scientific) gradually brought at -80ºC using a freezing container (Mr. Frosty™ Freezing Container, ThermoFisher Scientific) over 24h before being transferred to liquid nitrogen.

3.5. Virus

3.5.1. Isolates and reverse genetic viruses

Experiments were performed using isolates or reverse genetic versions of H3N8 equine influenza virus (EIVs). Full name, accession number of NS segment, type of virus, and abbreviation used in this work can be found at Table 3-7. For virus isolates, the mode of isolation, origin and potential passage(s) in chicken egg is unknown. Viruses were grown in MDCK cells at 37ºC, 5% CO_2 humidified atmosphere and virus stocks were grown at passage 2 in MDCK cells (MOI 0.1), and stored at -80ºC.

For infections, virus stocks were diluted in infection medium composed of DMEM, 0.3% Bovine Albumin Fraction V (BSA, 7.5% solution, ThermoFisher Scientific), 1% PS, and 1µg/ml of tosylsulfonyl phenylalanyl chloromethyl ketone (TPCK)-treated trypsin (Sigma-Aldrich) (Martinez-Sobrido and Garcia-Sastre, 2010).

Table 3-7: Viruses used in this work

Virus		Type of virus	Abbreviation
A/equine/Uruguay/1/1963	ACD85423	Isolate and Reverse genetic	U/63
A/equine/Miami/1/1963	ABY81497	Isolate	M/63
A/equine/SaoPaulo/1/1969	ACD85390	Isolate	SP/69
A/equine/Sussex/1/1989	ACD97430	Isolate	S/89
A/equine/Fontainebleau/1/1979	ACD85401	Isolate	F/79
A/equine/Kentucky/1/1991	ACA24672	Isolate	K/91
A/equine/LaPlata/1995	MF182460	Isolate	LP/95
A/equine/Kentucky/1995	MF182451	Isolate	K/95
A/equine/Kentucky/1999	MF182443	Isolate	K/99
A/equine/Kentucky/5/2002	ABA42429	Isolate	K/02
A/equine/Newmarket/5/2003	ACI48802	Isolate	N/03
A/equine/Ohio/1/2003	ABA42431	Isolate and Reverse genetic	O/03
A/equine/Mongolia/3/2013	MF182459	Isolate	M/13

The sequence of each NS segment, as well as NS1 coding sequence (CDS) and amino acid sequence, with corresponding accession numbers are available in Appendices 7 to 19.

3.5.2. Other viruses

Sendai virus (SeV), Cantell strain, was purchased from Charles River Laboratories and stored at -80ºC. The rVSV-GFP virus stock used in the bioassay assay was generated by transiently transfecting HEK293T cells in Opti-MEM with an expression plasmid encoding the Vesicular Stomatitis Virus (VSV) glycoprotein of surface (pVSV-G, kindly provided by Brian Willett) (Martinez-Sobrido et al., 2006) using TransIT-LT1 transfection reagent (Cambridge Biosciences) at 37ºC, 5% CO_2 for 36h (100mm sterile dish format, $5x10^6$ cells). At the end of the transfection period, the cells were infected with a previous stock of rVSV-GFP virus ($2.5x10^2$ TCID50) for 4h at 37ºC, 5% CO_2 humidified atmosphere in growth medium, then washed with PBS and maintained at 37ºC, 5% CO_2 humidified atmosphere in growth medium for a further 16h. The supernatant was then collected and filtered through a 0.45µm filter (Fisher Scientific), before being aliquoted and stored at -80ºC.

3.5.3. Virus rescue

Viruses were rescued using the eight-plasmid system of A/equine/Uruguay/1/1963 and A/equine/Ohio/1/2003 (Hoffman 2001). In this system, the cDNA of each eight influenza virus segments is inserted between the Polymerase I promoter (Pol_I h) and the Polymerase I terminator (t_i). This Pol_I transcription unit is flanked by the Polymerase II promoter of the human cytomegalovirus (Pol_{II} CMV) and the polyadenylation signal of the gene encoding bovine growth hormone (a_{II} BGH). After transfection of the eight plasmids in a co-culture (2:3) of HEK293T/MDCK cells (6-well plate format, $2x10^6$ cells/well), two types of molecules are synthesized: negative-sense vRNA by the cellular Pol_I, and transcription of mRNAs with 5'-cap structures and 3'-poly(A) tails by Pol_{II}. These mRNAs are translated into viral proteins. (Figure 3-1). Co-cultures were seeded 24h prior to viral rescue, and reverse genetic viruses are generated within 2–3 days following transfection. For transfection, cells were incubated in Opti-MEM (Opti-Minimum Essential Medium, Life Technologies) and transiently co-transfected using TransIT-LT1 (Mirus, Cambridge Bioscience), with 2.5µg of seven-ambisense O/03 or U/63 plasmids (PB2-, PB1-, PA-, HA-, NP-, NA-, M-pDP2002 plasmids) plus ambisense O/03 NS- or U/63 NS-pDP2002 plasmids or the NS mutant-pDP2002 constructs (O/03-K186E, O/03-230 and O/03-K186E-230; U/63-E186K, U/63-219, U/63-E186K-219, respectively). After 24h, the medium was replaced by infection medium. Virus-containing tissue culture supernatants (passage 0, P0) were collected 3 days post-transfection, clarified, and stored at -80ºC.

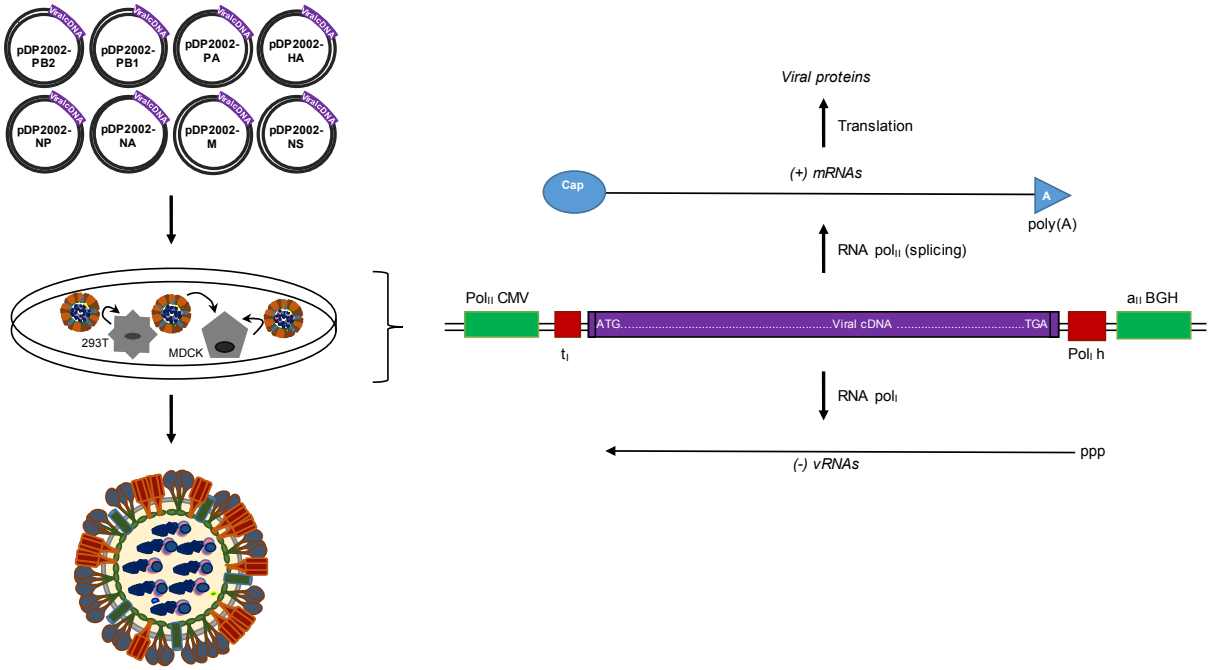


Figure 3-1: Eight-plasmid system for generation of Influenza A virus

Schematic representation of the IAV eight-plasmid rescue system. The left-hand side illustrates the co-culture system of HEK293T and MDCK cells. Cells are co-transfected with eight plasmids encoding for the eight segments of the IAV of interest. Reverse genetic viruses are generated within 2–3 days post transfection. On the right-hand side, a schematic representation of the polymerase (Pol) I–polymerase (Pol) II transcription system for synthesis of vRNA and mRNA is shown. The cDNA of each of the eight influenza virus segments is inserted between the Polymerase I promoter (Pol$_{I}$ h) and the Pol$_{I}$ terminator (t$_{I}$). This Pol$_{I}$ transcription unit is flanked by the Polymerase II promoter of the human cytomegalovirus (Pol$_{II}$ CMV) and the polyadenylation signal of the gene encoding bovine growth hormone (a$_{II}$ BGH). After transfection of the eight expression plasmids, two types of molecules are synthesized: negative-sense vRNA by the cellular Pol$_{I}$, and transcription of mRNAs with 5' cap structures and 3' poly(A) tails by Pol$_{II}$. These mRNAs are translated into viral proteins. Of note, the start codon (ATG) of each viral cDNA is directly following the Pol$_{II}$ transcription start site.

3.5.4. Creating virus stocks

For experimental infection, a minimum of two viral stocks of each virus was rescued and grown independently. P0 stocks were used to infect fresh MDCK cells to generate a passage 1 (P1) stock. After 2 to 3 days of infection (when 80% of the monolayer was destroyed), supernatant was collected, cleared and stored at -80°C. Viral titres of P1 stocks were determined as described below. The P1 stocks were then used to generate passage 2 (P2) viral stocks. To this end, MDCK cells were infected with P1 stocks at MOI 0.01 and after 2 to 3 days of infection (when 80% of the monolayer was destroyed), supernatants were collected, cleared and stored at -80°C. The sequence of each genomic segment was checked by Sanger sequencing and the size of each segment was checked by PCR and 1% agarose gel electrophoresis (Appendices 20 to 31). Viral titres of P2 stocks were determined as described

below. For experimental infections, a minimum of two viral stocks for each virus were rescued and grown independently.

3.5.5. Experimental viral infections

Confluent monolayers of MDCK cells (12-well plate format, triplicates, $5x10^5$ cells/well) or E.Derm cells (12-well plate format, triplicates, $2.5x10^5$ cells/wells) were infected (MOI 0.01 and 0.1, respectively) with the indicated viruses and placed at 37ºC, 5% CO_2. After 1h incubation, cells were washed with PBS and infection medium was replaced with 500µl of fresh growth medium. Tissue culture supernatants were collected at various times pi and stored at -80ºC, and cells were fixed in 0.1% buffered formalin at 4ºC for 16h and kept for flow cytometry analysis. Each experiment was repeated three times independently. Viral titres were determined by immunofocus assay in MDCK cells. Titrations were repeated three times independently, and the mean value and standard error mean were calculated using GraphPad Prism7 (GraphPad Software Inc. San Diego, CA, USA).
Of note, the presence of deficient interfering particles in viral stocks was not assessed. In addition, the level infectious particles versus genomic information was not compared between viruses. Thus, the differences observed throughout this work could either reflect differences due to the mutations introduced in NS1 of /03 and U/63 or NS segment swap between O/03 and U/63, or be due to differences in infectivity levels and ability to induce an IFN response between WT and mutant viruses.

3.5.6. Determination of viral titre and plaque phenotype by immunofocus assay

Viral titres were determined by immunofocus assay (focus forming units, FFU/ml) in MDCK cells. Confluent monolayers of MDCK cells (48-well plate format, triplicates, $1.25x10^3$ cells/well) were infected with serial dilution (1:10) of the viral stock of interest and placed at 37ºC, 5% CO_2. Plates were gently rocked every 10 minutes. After 1h incubation, cells were washed with PBS and infection medium was replaced with a 50:50 2.4% Avicell:2X MEM overlay for 48h. 2.4% Avicell solution was prepared as follow: 2.4% (w/v) of Avicell (Avicell RC/CL, Microcrystalline cellulose & Sodium carboxymethylcellulose1, Sigma-Aldrich) in 100ml of H_2O. Autoclaved and stored at room temperature until use. At 48hpi, the overlay was discarded, cells were washed 3 times with PBS and fixed with 80% ice-cold acetone solution (Acetone AR, >99.5%, Sigma-Aldrich) for 10 minutes at room temperature. The

plates were let to dry overnight at room temperature. then treated with 1% Triton X-100 PBS solution (Triton™ X-100, Sigma-Aldrich-Aldrich) for 10 minutes at room temperature, followed by 1h of incubation with 10% NGS PBS solution at room temperature. This was followed by 3h of immunoblotting at room temperature in 10% Normal Goat Serum plus PBS with a monoclonal anti-influenza A virus nucleoprotein (NP) antibody (clone HB65, European Veterinary Laboratory) (Table 3-8). After a 3-step washing with PBS, a horseradish peroxidase-conjugated rabbit anti-mouse IgG antibody (AbD Serotec, UK) (Table 3-8) was used in PBS solution for a further 1h at room temperature. A color development method was used to reveal the immunofocus using the TrueBlue peroxidase substrate (Insight Biotechnology). 50µl of substrate per well was used under agitation for 10 minutes of incubation, and stop with water. Viral titres were calculated by counting the number of blue plaques and were expressed as log10 PFU per millilitre.

For plaque phenotype, a confluent monolayer of MDCK cells (6-well plate format, triplicates, $6.4x10^4$ cells/well) were infected with serial dilution (1:2) of the viral stock of interest and placed at incubation at 37ºC, 5% CO_2, humidified atmosphere. The same protocol as described above was then followed. For colour development, 2ml of TrueBlue peroxidase substrate was used per well.

3.5.7. Viral protein staining for flow cytometry and confocal microscopy

Cells were permeabilized with 1% Triton X-100 for 10 minutes, and blocked in PBS 10% Normal Goat Serum (Gibco, Life Technologies) for 1h. Cells were then incubated with anti-NP antibody, rabbit polyclonal anti-NS1 protein antibody (Genscript) or rabbit STAT1 polyclonal antibody (Santa Cruz) overnight at 4ºC. Cells were then washed twice with PBS and incubated for 4h with a rabbit anti-mouse IgG Alexa fluor 488 (Cell signalling) or donkey anti-rabbit IgG Alexa fluor 555 (Cell signalling), before analysis by flow cytometry (Guava Flow Cytometer, Merck) or fixed in VECTASHIELD Mounting Medium with DAPI (VECTASHIELD® Hard+Set™ Mounting Medium with 1.5µg/ml DAPI, Vector laboratories) and analysed by confocal microscopy.

3.5.8. Viral growth kinetics with Ruxolitinib or universal IFN

Ruxolitinib (Selleck Chemicals), a JAK1/2 inhibitor (Stewart et al., 2014), was prepared as 10mM stocks in dimethyl sulfoxide (DMSO). Several doses of Ruxolitinib were tested in the presence of 500 Units (U) of universal type I IFN alpha (uIFN, pbl Assay Science)

(Figure 3-2). To measure virus growth kinetics in cells rendered un-responsive to type I IFN or in the presence of exogenous type I IFN, E.Derm cells were treated with Ruxolitinib at a concentration of 4μM as determined in or with 500U of uIFN. Treatments were started 24h prior to infection and maintained at the same concentration for the whole experiment.

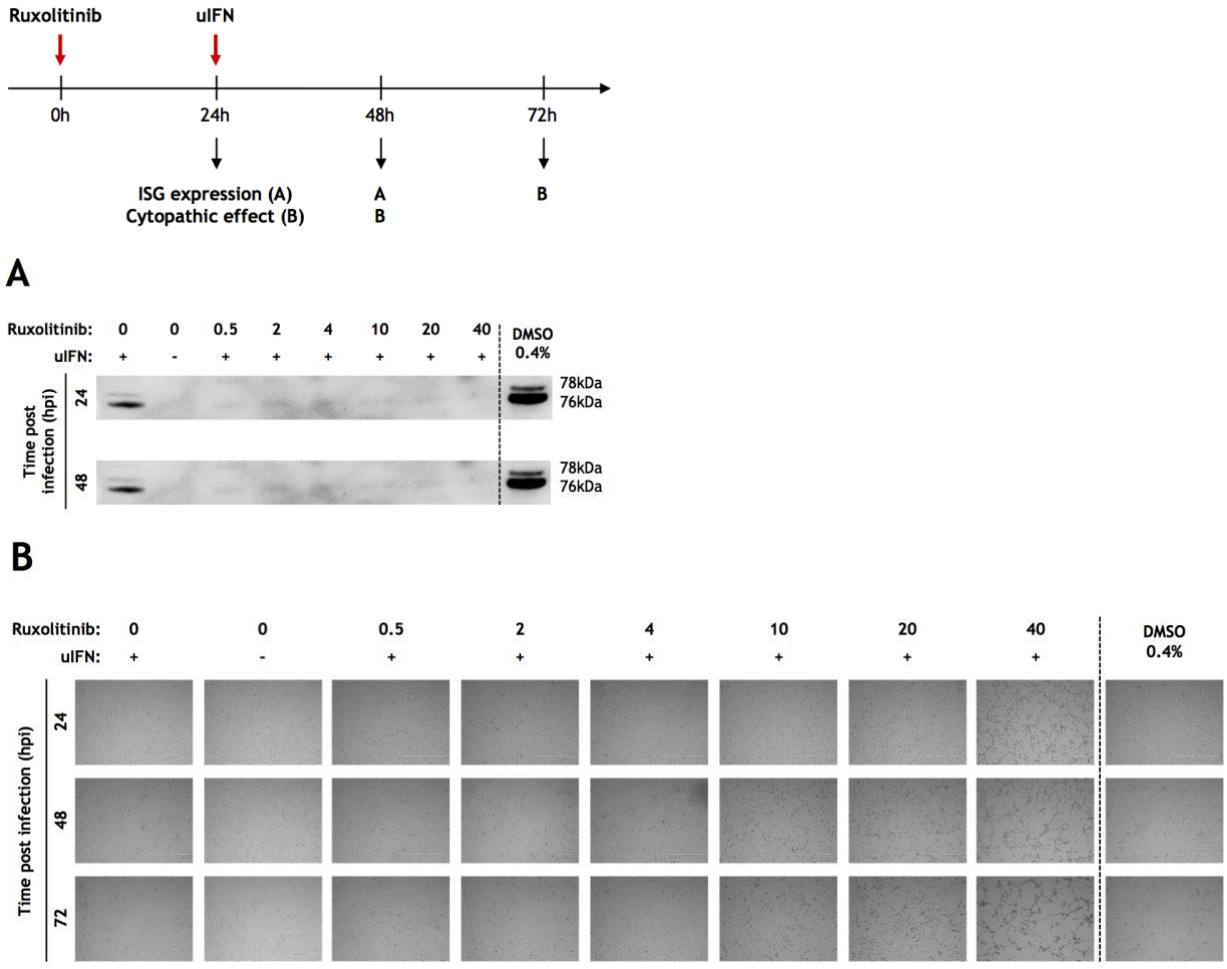


Figure 3-2: Test of Ruxolitinib concentrations

E.Derm cells were treated with different amounts of Ruxolitinib (0.5 to 40μM) or mock treated with 0.4% of DMSO for 24 h prior to treatment with universal IFN (uIFN). At the indicated time post infection (hpi) **(A)** Mx proteins expression was assessed by Western blot (78 and 76 kDa), and **(B)** cytopathic effect was evaluated using a confocal microscope (transmitted light mode).

3.6. Antiviral cytokine production and general protein shutdown.

E.Derm cells (12-well plate format, triplicates, $2.5x10^5$ cells/well) were infected (MOI 0.1) with the indicated viruses for a total of 72h. At the indicated times post-infection, supernatants were collected and stored at -80°C for further analysis by bioassay, while cells

were treated for 1h at 37ºC, 5% CO_2 with Puromycin (20µg/ml in DMEM, 15% FBS) prior to lysis in protein disruption buffer + β-mercaptoethanol and stored at -80ºC for further analysis by western blot. Puromycin is a well-known antibiotic that competes against aminoacyl tRNA on the ribosome A site (Monro et al., 1968). As such, puromycin enables examination of total protein production without requiring transfection, radio-labeling, or the prior choice of a candidate gene (Starck et al., 2004). If cells are incubated with puromycin, lysed and immunoblotted using an anti-puromycin antibody, all the proteins being produced will be immunostained as puromycin will be incorporated at the C-terminus of all nascent proteins. For the IFN bioassay, supernatants were UV-inactivated for 5 minutes at room temperature and used to treat fresh E.Derm cells (48-well plate format, $6x10^4$ cells/well, triplicates) for 24h. The cells were then infected with rVSV-GFP virus ($2.5x10^2$ TCID50) for 8h, then trypsinized and fixed in 0.1% buffered formalin for 16h at 4ºC. The percentage of GFP-expressing cells were then analysed by flow cytometry. For controls, E.Derm cells were mock treated or treated with 500U of uIFN. GFP expression of mock-treated cells infected with rVSV-GFP was considered as 100% and GFP expression of uIFN-treated cells infected with rVSV-GFP was considered as 0%. Mean values and standard error means were calculated with GraphPad Prism7.

3.7. SDS-PAGE & Western blot analysis

Cells were lysed in protein disruption buffer (0.125 M Tris-HCl (pH 6.8), 4% (w/v) SDS, 25% (v/v) glycerol, 0.02% (w/v) bromophenol blue) + β-mercaptoethanol and stored immediately at -80ºC. Samples were boiled for 15 minutes at 95ºC prior to polypeptide separation by SDS-PAGE on NuPAGE Novex 4-12% Bis-Tris protein gels (ThermoFisher Scientific) using 20X NuPAGE MES SDS Running Buffer (Invitrogen), diluted in H2O. Proteins were detected by Western blotting following to transfer to nitrocellulose membranes using 20X NuPAGE Transfer Buffer (Novex, Life technologies) diluted in H_2O. The membranes were blocked for 1h at room temperature in 5% blotting-grade blocker (nonfat dry milk, Bio-Rad) TBS 0.1% Tween-20 (TWEEN® 20, Sigma-Aldrich) and immunoblotted overnight at 4ºC in 5% blotting-grade blocker TBS-0.1% Tween-20 with the antibody of interest. 10X TBS buffer (Tris base 48.4g, NaCl 160g, H_2O up to 2L, pH adjusted at 7.6, Sigma-Aldrich) was diluted to 1X in H_2O for use. Antibodies used in this study are summarized in table 3-8. The chemiluminescent signal was detected using Amersham ECL Prime Western Blotting Detection Reagent (GE-Healthcare) and captured with ChemiDoc XRS+ System (Biorad).

3.8. RNA sequencing

Confluent monolayers of E.Derm cells (12-well plates, triplicates, $2.5x10^5$ cells/wells) were infected with O/03 and revertant viruses (MOI 0.1 to 1) or mock infected at least three times independently. At 8 hpi cells were washed with PBS, immunostained with the anti-NP antibody and the proportion of infected cells was determined by flow cytometry. The samples containing a similar proportion of infected cells were selected for transcriptomic analysis, and cells were lysed with 500µl of TRIzol (ThermoFisher Scientific) for further RNA extraction. Total RNA was extracted using the TRIzol method and further purified using the RNeasy mini spin columns (Qiagen), including an on-column DNase I digestion step (Qiagen) according to the manufacturer's protocol. RNA concentration was measure with Qubit and the Qubit RNA HS Assay Kit (ThermoFisher Scientific) following the manufacturer's protocol. The ribosomal (r)RNA integrity number was measured using an Agilent 2100 BioAnalyzer (Agilent Technologies). Library preparation and sequencing were carried out by the Viral Genomics and Bioinformatic team of the MRC University of Glasgow Centre for Virus Research. 4.5ug of total RNA was enriched by selectively depleting rRNA using the RiboMinus™ Eukaryote Kit v2 (Ambion, Life Technologies). The sequence reads were processed according to the Tuxedo pipeline (Trapnell et al., 2013). Read quality was assessed using FastQC, and TopHat2 and Bowtie2 were used to map short reads against the Equus caballus 2 genome (GCA_000002305.1). A list of differentially expressed genes (DEGs) to mock-infected samples was generated using CuffDiff2 (genes with Benjamini Hochberg-p value <0.05 were considered significant) (Ratinier et al., 2016). A complete list of DEGs in all samples can be found in Appendices 35 to 67.

3.9. Analysis of EIV NS1 amino acid sequences

For the phylogenetic analysis of EIV NS1, a total of 170 NS sequences were collected from the NCBI Influenza Virus Resource database. Four NS segments sequenced for this study were also added (42). SeaView (Version 4.6.1) (43) was used to align the NS1 coding regions and the final alignment (http://datadryad.org/review?doi=doi:10.5061/dryad.673bj) was edited manually to remove NEP coding regions. BEAST (Version 1.8.4) (44) was used to infer maximum clade credibility trees, as well as a strict molecular clock and HKY85+G model of nucleotide substitution. Each codon position was estimated with separate substitution rates and nucleotide frequencies. Two individual chains were run until convergence was achieved. To analyse NS1 single nucleotide polymorphisms (SNP), the 175 NS1 amino acid sequences were aligned with CLC genomic workbench 8 (QIAGEN Bioinformatics) (Appendices 1 and 2)

and the frequency of each amino acid found at each of the 230 composing the NS1 protein was recorded. A threshold of 7% was used in accordance with the "best practices for evaluating single nucleotide variant calling" (Olson Frontiers in Genetics 2015).

3.10. Graphing and statistical analysis

All statistical analyses were conducted using GaphPad Prism software. Unless stated otherwise, significance was calculated by Two-way ANOVA and a Bonferroni's multiple comparisons post hoc test.

3.11. Antibodies

All primary and secondary antibodies used in this work can be found in Table 3.8.

Table 3-8. Antibodies

Target	Antibody	Dilution
NS1	Polyclonal rabbit, IgG, (GenScript)	1:1000 WB 1:500 IF
NP	Monoclonal mouse, IgG2α, clone HB65 (European Veterinary Laboratory)	1:1000 WB 1:1000 IF
MX1	Monoclonal mouse, IgG2α, clone M143 (kindly provided by Georg Kochs)	1:500
ISG15	Polyclonal rabbit, IgG, fusion protein ag8782 (Proteintech)	1:500
STAT1 (p84/p91)	Polyclonal rabbit, IgG, clone M-22 (Santa Cruz)	1:1000
Caspase 3	Polyclonal rabbit, IgG (Cell signaling)	1:1000
Cleaved caspase 3	Polyclonal rabbit, IgG, clone Asp175 (Cell signaling)	1:500
HA-tag	Rabbit polyclonal, IgG (Sigma-Aldrich)	1:1000
Puromycin	Monoclonal mouse, IgG2ακ, clone 12D10 (Millipore)	1:1000
γ-Tubulin	Polyclonal rabbit, IgG, clone AK-15 (Sigma-Aldrich)	1:2000
Rabbit	Polyclonal donkey, IgG, HRP-conjugate (GE-healthcare)	1:2000
Mouse	Polyclonal rabbit, IgG, HRP-conjugated (Starlab, Biorad)	1:2000
Mouse	Polyclonal donkey, IgG, Alexa Fluor® 488-linked (ThermoFisher Scientific)	1:2000
Rabbit	Polyclonal mouse, IgG, Alexa Fluor® 647-linked (ThermoFisher Scientific)	1:2000

Chapter 4

Evolution of NS1 during EIV post-transfer adaptation

4. Evolution of NS1 during EIV post-transfer adaptation

4.1. Introduction

The principal natural reservoir of IAVs is wild aquatic birds, and these viruses also circulate in some mammalian host populations, including horses (Parrish et al., 2015, Donatelli et al., 2017). Interspecies transmission events, notably from bird to mammal occur periodically, but most infections result in spillover events or isolated outbreaks, and very few lead to the establishment of novel endemic lineages (Parrish et al., 2015, Webby and Webster, 2001). In this work, H3N8 EIV was used as a model to study the adaptation process of an avian-origin IAV to a mammal. EIV was first reported in January 1963 in Florida (USA) from horses recently imported from Argentina (Scholtens et al., 1964, Waddell et al., 1963), and within two years following emergence, this virus was responsible for a major transcontinental pandemic (Paccaud and Paccaud, 1967). This virus has been circulating since then in the equine population, causing major outbreaks of disease worldwide (Burrows et al., 1981, Cowled et al., 2009, Livesay et al., 1993, Ito et al., 2008, Newton et al., 2006, Virmani et al., 2010, Alves Beuttemmuller et al., 2016, Perglione et al., 2016, Yamanaka et al., 2008).

To establish a new lineage in a mammal, an avian IAV has to overcome multiple barriers, such as the immune system of the new host or molecular incompatibilities between the host cellular machinery and the viral proteins (Long et al., 2016). The occurrence of multiple advantageous mutations might have been important for the virus to successfully overcome species boundaries and adapt to the new host (Duffy et al., 2008). The role of HA mutations in virus-host interaction has been one of the main focus of the scientific community for decades (Lipsitch et al., 2016). Mutations in other viral segments such as PB2, NP, NA, M, and NS also occur (Le et al., 2009, Sakabe et al., 2011, Dankar et al., 2013), but their contribution to IAV adaptation to a new host is still incompletely understood.

The evolution of the EIV lineage has resulted in accumulation of mutations in each of the viral genomic segments (Murcia et al., 2010), and the lineage has been divided in different clades (Daly et al., 1996, Murcia et al., 2011b, Lai et al., 2001). Since the late 1990's, viruses that have caused disease in horses belonged to the Florida clade 1 (FC1) or clade 2 (FC2) (Bryant et al., 2009, Lewis et al., 2011), which are endemic in North America, or Europe and Asia, respectively (Yamanaka et al., 2008, Cowled et al., 2009, Qi et al., 2010, Virmani et al., 2010, Gildea et al., 2013, Yondon et al., 2013, Woodward et

al., 2014, Gagnon et al., 2007, Damiani et al., 2008, Landolt, 2014). The Florida clades are notably distinguished from other clades by several amino acid changes in NS1, such as residue 44, 59, 71, 86, 216, and by an 11-amino acid C-terminal truncation of the protein (Barba and Daly, 2016).

The NS1 protein is encoded on a collinear mRNA derived from the NS segment, which also encodes for the nuclear export protein (NEP) upon splicing (Inglis et al., 1979, Lamb and Choppin, 1979). NS1 and NEP share their first 10 amino acids, and NS1 bases 503-690 overlap with NEP bases 31-221, although on a shifted reading frame (Lamb and Lai, 1980). NS1 is a major virulence factor common to all IAVs, whose best described function is to antagonise the IFN pathway (Garcia-Sastre et al., 1998, Kochs et al., 2007a, Egorov et al., 1998), an essential component of the host immune response to viral infection.

NS1 can act both at pre-transcriptional and co-/post-transcriptional levels, although variations exist between viral strains (Geiss et al., 2002, Kochs et al., 2007a, Twu et al., 2007, Hayman et al., 2007).

NS1 pre-transcriptional activities involve the interference with the RIG-I pathway, mainly by associating with two positive regulators: the ubiquitin ligases TRIM25 (Gack et al., 2009) and RIPLET (Rajsbaum et al., 2012). NS1 residues R38, K41, E96, E97 seem to play a central role in this function (Wang et al., 1999, Talon et al., 2000) (Pichlmair et al., 2006, Gack et al., 2009, Rajsbaum et al., 2012). Residue 196 has also been shown to regulate NS1 antagonism of IFN-β induction (Kuo et al., 2010) through an IRF3-dependent mechanism (Talon et al., 2000, Krug, 2015).

While the interaction of NS1 with the components of the RIG-I signaling axis are conserved among virus isolates, the extent and efficacy of the inhibitory effect varies significantly among strains (Hayman et al., 2006, Kuo et al., 2010). The same applies for NS1 interference with IRF3 function (Kuo et al., 2010).

NS1 co-/post-transcriptional activities comprise the prevention of nuclear processing of RNA polymerase II transcripts (Twu et al., 2007, Hayman et al., 2006, Hayman et al., 2007, Kochs et al., 2007b, Noah et al., 2003). In eukaryotic cells, the 3'-end of primary transcripts are cleaved upon recognition of a conserved 'AAUAAA' sequence some ten to thirty bases upstream of the cleavage site. The Cleavage and Polyadenylation Specificity Factor (CPSF) is a polyprotein complex formed by four subunits that recognizes the 'AAUAAA' sequence, binds to the nascent mRNA and catalyzes the endo-nucleolytic step and the subsequent addition of a poly(A) tail. The smallest component of the CPSF

complex is a 30 kDa protein (CPSF30), which possess five C3H zinc fingers (Tian and Manley, 2017). NS1 can bind CPSF30 through two of these zinc fingers (F2F3), and in doing so it inhibits the binding of the whole CPSF complex to pre-mRNA, thus preventing cleavage and polyadenylation of cellular pre-mRNA (Das et al., 2008). Of note, viral transcripts are not affected by the repression of CPSF, because their polyadenylation is directly catalyzed by the viral polymerase (Ayllon and Garcia-Sastre, 2015). The CPSF30 binding site of NS1 is composed of a patch of highly conserved hydrophobic residues centered around amino acid 186 (184 to 189) (DeDiego et al., 2016), and these amino acids embed into a pocket on the F2F3 domain of CPSF30 (Hale et al., 2010b). Other amino acids have been described to stabilize NS1-CPSF30 interaction, such as G96, F103, M106, K108R, D125E, although they are not strictly part of the binding interface (Hale et al., 2010b, Kochs et al., 2007b, Das et al., 2008, Shimizu et al., 1999). Interestingly, variations in CPSF30 binding seem to arise commonly upon viral adaptation to different hosts or during laboratory passage (Kochs et al., 2007a, Twu et al., 2007, Hayman et al., 2007). Indeed, mouse-, egg- and swine-adapted viruses generally have NS1 proteins that bind weakly to CPSF30, whereas those from most human-circulating strains are strong binders (Hossain et al., 2008, Brown et al., 2001, Elderfield et al., 2014, Hale et al., 2010b). However, to date there is no clear evolutionary advantage associated with NS1-CPSF30 interaction.

NS1 has also been shown to interfere with the function of another protein involved in the processing of pre-mRNA, the poly(A)-binding protein II (PABPII) through residues 223-237 (Li et al., 2001). Although, the exact mechanism is still unclear, NS1 would further reinforce viral shutoff of host gene expression during infection by binding to PABPII (Chen et al., 1999).

Interestingly, successive truncation/elongation events in the NS1 C-terminus have been described during evolution of different IAV subtypes, including EIV (Suarez and Perdue, 1998, Dundon and Capua, 2009). However, to date no clear evidence of evolutionary advantage has been associated with these events. The C-terminal tail of NS1 also accounts for other functional differences between viral strains. Indeed, based on bioinformatics and sequence analysis, residues 227 to 230 in the C-terminus of many 230-amino acid long NS1 proteins possess a PDZ binding motif (PBM) (Obenauer et al., 2006). PDZ domains are protein-protein recognition modules within a multitude of proteins that organize diverse cell-signaling assemblies (Javier and Rice, 2011). The most common PBMs found in avian IAV NS1s is composed of 'ESEV/EPEV', which have been found to interact with PDZ domain-containing proteins. This implies that when present in mammalian cells,

avian NS1 proteins can interact with PDZ domain-containing proteins and disrupt important cellular pathways (Obenauer et al., 2006). The majority of EIVs in circulation until the late 1990s, expressed a full length NS1 (230 amino acids) possessing an avian PBM 'ESEV/EPEV' at residue 227-230. This motif was lost after the late 1990s when the protein got truncated of eleven amino acids at its C-terminus (Murcia et al., 2011b, Barba and Daly, 2016). The loss of the ESEV/EPEV motif likely affected NS1's ability to interact with PABPII or PDZ-containing proteins, however to date, no clear evolutionary advantage has been associated with this change.

Many caveats in our understanding of the NS1 protein function remain, with notably little knowledge on the impact of NS1 amino acid changes on viral evolution and adaptation to horses. In this chapter, a phylogenetic approach was associated with cloning techniques and *in vitro* assays to identify genetic markers of adaptation of EIV to horses in the NS1 protein.

Aims: The aims of this chapter were to identify molecular determinants of EIV NS1 evolution; and to analyze the role of specific mutations on NS1 function at the protein level.

4.2. Results

4.2.1. Evolution of NS1 amino acid sequence from 1963 to 2014

To assess EIV NS1 phylogeny, 175 full length NS sequences of equine H3N8 IAVs were used. A maximum clade credibility tree was generated following the methodology described in Chapter 3, section 3.10. As shown in Figure 4-1, several clades in the NS1 phylogenetic tree could be distinguished. Following the nomenclature found in previous work (Murcia et al., 2011a), the different clades were named in a chronological fashion from clade I to X. The so-called clades IV and VI mentioned in Murcia and colleagues' work were not distinguishable in this analysis, however, for consistency viruses from these clades were referred to as 'clade IV-VI'. Each of these clades included a number of viruses that shared a common ancestor supported by high bootstrap values. Clade III and IX contained viruses from Europe, clade II comprised viruses isolated from Japan and Algiers, clade I and VIII included viruses from South America, and clades IV-VI, V and VII were composed of viruses isolated from North America (Georgia & Kentucky, Kentucky, and

California, respectively). Two major bifurcation events could be observed along the main trunk of the tree (late 1980s and 1990s), and during the second bifurcation the Florida clades 1 and 2 (FC1 and FC2) emerged. As mentioned previously, EIVs in circulation in the equine population today belong to FC1 or FC2, and are mainly isolated from North America, or Asia and Europe, respectively.

The 175 NS1 amino acid sequences were then aligned and the frequency of amino acid found at each of the 230 residues composing the NS1 protein was recorded. To evaluate single nucleotide polymorphisms (SNP) variants with a viral population, a threshold of 5% is generally applied (The 1000 genomes consortium Nature 2010). However, Olson and colleagues published in 2015 the "best practices for evaluating single nucleotide variant calling" in which they showed that errors associated with Illumina sequencing can be as high as 6% (Olson Frontiers in Genetics 2015), thus, placing a SNP threshold at 7% would reduce the risk of false positives when defining a SNP of high frequency. Using this threshold, 149 residues were found completely conserved among NS1 proteins, while 66 residues were changed throughout evolution, but were considered as punctual SNP as the number of isolates carrying these mutations remained below the 7% threshold (Table 4-1). More importantly, 15 residues changed at high frequency (>7%) were found and most of these changes were fixed at the population level.

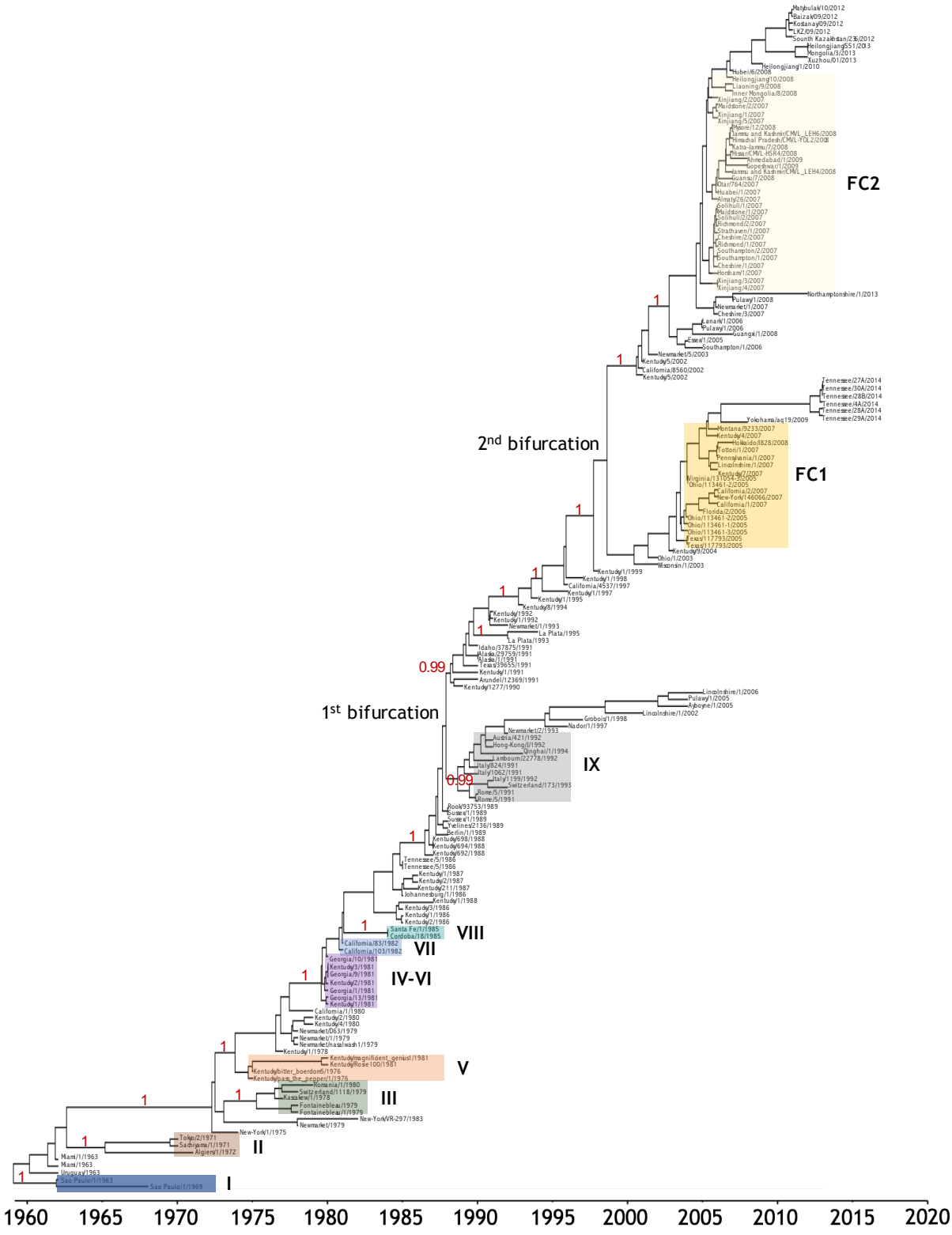


Figure 4-1: Phylogenetic relationship of EIV NS1 genes from 1963 to 2014.

A total of 175 EIV NS1 sequences were used to infer a maximum clade credibility tree. Horizontal branches are drawn to a scale of nucleotide substitutions per site, and the tree is rooted on A/equine/Uruguay/1/1963. Bootstrap values above 0.99 are shown in red. Colored boxes represent distinct clades of EIV numbered from I to IX in a chronological fashion. The two major bifurcations of the main trunk are indicated in black. The 2nd bifurcation in the late 1990s gave rise to the Florida clade 1 (FC1) and clade 2 (FC2).

Table 4-1: Residue changes in EIV NS1 from 1963 to 2014.

Residue evolution	Number of residues	Position & type of amino acid found				
Conserved	149	1 M 2 D 4 N 6 V 8 S 9 F 10 Q 11 V 12 D 13 C 14 F 15 L 16 W 17 H 20 K 21 R 22 F 23 A 25 Q 26 E 27 L 29 D 30 A 31 P 32 F 33 L 34 D 35 R 36 L 37 R	38 R 39 D 40 Q 41 K 42 S 43 L 46 R 50 L 54 I 55 E 57 A 60 A 61 G 63 Q 65 V 68 I 73 S 75 E 78 K 80 T 82 A 83 S 85 P 88 R 89 Y 90 L 92 D 93 M 94 T 97 E	99 S 100 R 102 W 103 F 105 L 106 M 107 P 108 K 109 Q 110 K 113 G 114 S 115 L 116 C 118 R 119 M 120 D 121 Q 123 I 125 D 126 K 127 N 128 I 130 L 131 K 132 A 133 N 134 F 135 S 136 V	137 I 141 L 142 E 143 T 145 I 147 L 148 R 149 A 150 F 151 T 152 E 153 E 154 G 157 V 158 G 159 E 160 I 161 S 162 P 163 L 164 P 165 S 167 P 168 G 169 H 170 T 171 N 172 E 173 D 174 V	175 K 176 N 177 A 178 I 182 I 183 G 184 G 187 W 188 N 189 D 190 N 193 R 195 S 196 E 198 L 199 Q 200 R 201 F 203 W 208 E 211 R 214 F 215 P 217 K 219 K *221 K* *222 M* *225 T* *226 I*
Substitution at low frequency	66	3 S PF 5 T P 7 S L 18 V I 19 R C 24 D N 28 G S 45 G R 47 G S 49 T A 51 G AD 52 L MQ 53 D GN 56 T I	58 T AI 62 K R 64 I M 66 E GK 69 L M 70 E K 72 E K 74 D N 76 A E 77 L FP 79 M R 81 I T 87 S P 91 T I	95 L F 96 D EG 98 M I 101 D E 104 M T 111 V I 117 I V 122 A T 124 M V 129 I V 138 F L 139 E KA 140 R MKT 144 L V	146 L I 155 A TS 156 V I 166 L F 179 G ER 180 V I 181 L I 185 L F 191 T K 192 V I 197 T AI 202 A T 204 R K 205 S N	206 S C 207 H NY 209 N ID 210 G R 212 P S 213 S T 218 Q * *224 R K* *227 E K* *229 E K*
Substitution at high frequency	15	44 K<>R E 48 S>I GN 59 R>H	67 R>Q G 71 E>K 84 V>I	86 A>T/V 112 A>T I 186 E>K	194 V>I 216 P>S 220 R>* Q	*223 A>E* *228 S>P* *230 V>I*

The first column divides into 3 categories the NS1 residues found in the EIV population: conserved, substitution at low and high frequency. The second column indicates the total number of residues found in each category. The third column details the nature of the amino acid found at each residue. Residues marked in black are 100% conserved in all EIV NS1 sequences studied. Green residues indicate the main amino acid found at each specific position. Small grey residues indicate variants present at low frequency in NS1 in the EIV population. Red residues indicate the amino acid substitution found in viruses that follow the main trunk of the NS1 phylogenetic tree. *, indicates a stop codon; >, indicates a substitution fixed throughout evolution; <>, indicates an amino acid reversion. Of note, residues in NS1 C-terminal tail that are not present in the short version of NS1 (221 to 230) are represented in *italic*.

Of the 15 high frequency SNPs, 14 coded for non-synonymous substitutions: K44R, S48I, R59H, R67Q, E71K, V84I, A86T/V, A112T, E186K, V194I, P216S, A223E, S228P, V230I, and one for a non-sense mutation: R220*. The introduction of this premature stop codon was responsible for a 11-amino acid C-terminal truncation of the protein, as previously described (Quinlivan et al., 2005). Of note, only one substitution (K44R) reverted to the original amino acid, while the 14 others were maintained at the population level.

A schematic representation of these 15 residue changes were then represented along a time scale in order to compare their respective introduction during NS1 evolutionary history (Figure 4-2). It became apparent that some of these changes occurred at the same period. In example, A112T and E186K substitutions appeared both in the early 1970s, while R44K reversion occurred at the same period than R59H, E71K, A86T and V230I substitutions (early 1990s). This last event was closely followed by P216S substitution and NS1 C-terminal truncation (R220*) in the late 1990s. Interestingly, some clades that separated from the main trunk of the tree did not introduce some of these residue changes. This was notably the case for clade IX, which maintained R59, E71, V194, P216, R220, V230 in the 1990s, or for clade III, which kept S228 in the mid-1970s. In both cases, the clades died out shortly after separation from the main trunk (early 1980s for clade III and mid-2000s for clade IX).

In addition, an arginine (R) could be found at position 67 in the main trunk until the 1980s when it changed for a glutamine (Q). Q67 was already found at earlier times in minor clades, i.e. clade I and V, in the early 1960s and mid-1970s, respectively. Surprisingly, these two clades disappeared only a few years after circulation (late 1960s and early 1980s, respectively).

Furthermore, two substitutions only occurred in one of the currently circulating Florida clades, S48I and V84I. Indeed, NS1 of FC1 viruses maintained S48 and V84, while NS1 of FC2 viruses introduced I48 and I84 in the early 2000s.

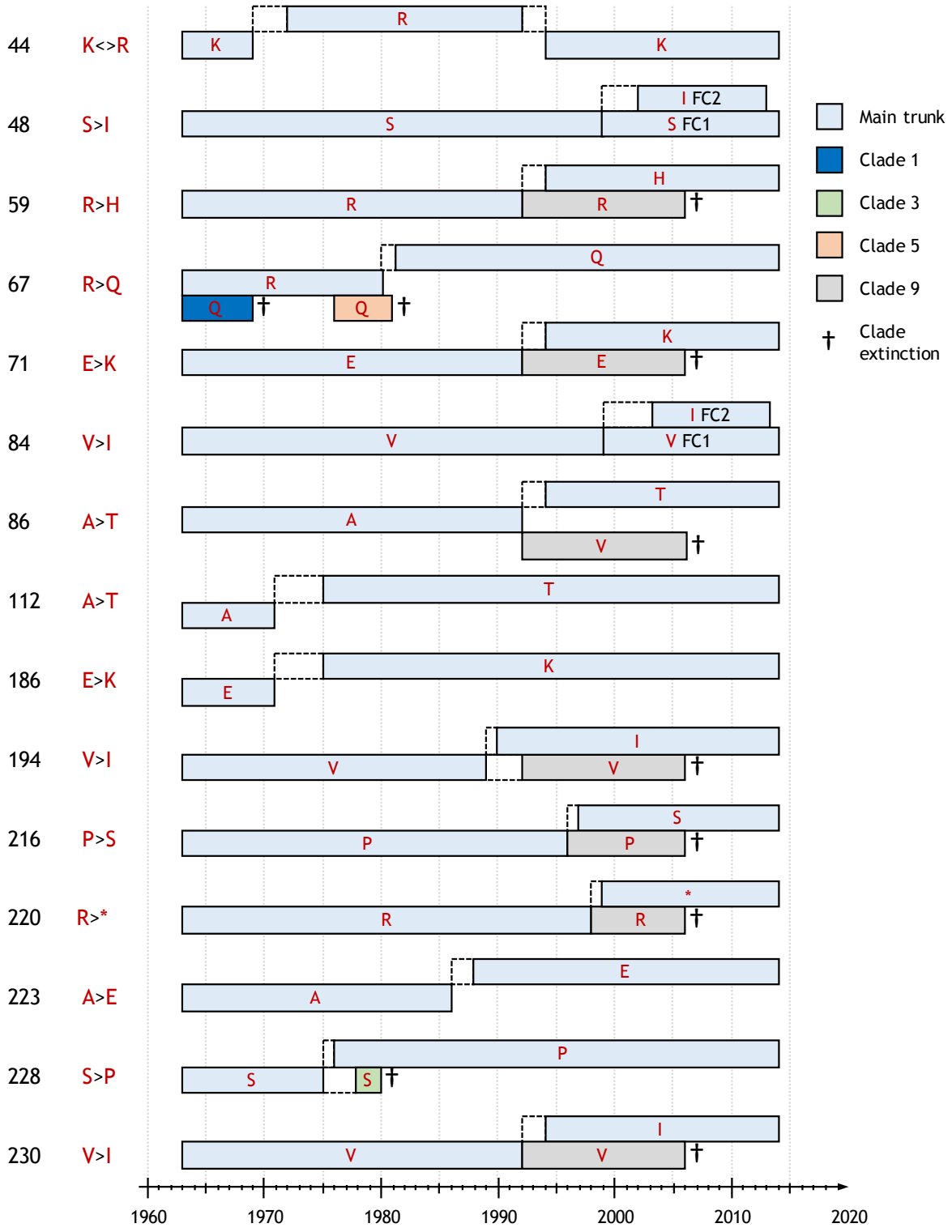


Figure 4-2: Schematic representation of NS1 high frequency residue changes from 1963 to 2014.

Red residues indicate amino acid substitutions found in viruses that follow the main trunk of the NS1 phylogenetic tree (Figure 4-1). A change of direction of the main trunk (light blue) indicates the estimated date of apparition a substitution. Dotted lines represent a period during which the nature of the residue found at the corresponding position is unknown. Residues found in other clades for the indicated positions are shown. The estimated time of extinction of each clade are indicated.

When looking for amino acids important for NS1/CPSF30 interaction (D96, F103, M106, K108, D125, L144, and 184-GLEWND-189) (Hale et al., 2010b, Kochs et al., 2007b, Das et al., 2008) it appeared that only residue 186 was changed at high frequency throughout evolution (Table 4-1). Of note, L144, L185 and D96, the latter being also involved in TRIM25 inhibition (Gack et al., 2009) were changed in a limited number of isolates, while all others were 100% conserved (Table 4-1 & Appendices 1&2).

Furthermore, among residues important for NS1 dsRNA binding and inhibition of IFN induction (via IRF3 or TRIM25) (Wang et al., 1999, Talon et al., 2000, Pichlmair et al., 2006, Gack et al., 2009, Rajsbaum et al., 2012, Kuo et al., 2010) (Liu et al., 1997, Chien et al., 1997), the majority were conserved at 100% (P31, D34, R35, R38, K41, R46, E97, E196), while others changed in a small number of isolates (T5, G45, T49).

Most residues important for NS1 subcellular localization were also conserved. Indeed, all residues important for the protein nuclear import (residues 35 to 41) were 100% maintained (Greenspan et al., 1988; Li et al., 1998), while residues important for nuclear export (residues 138 to 147) (Tynell et al., 2014) were mostly conserved (141 to 143 and 145, 147), with a few exceptions that changed at low frequency in the population (138 to 140 and 144).

Additionally, Y89 important for PI3K activation (Hale et al. 2006, 2010) was 100% conserved, and residues important for PKR inhibition (123 to 127) (Li et al., 2006, Min et al., 2007, Tan & Katze, 1998) were mostly conserved, with the exception of residue 124, which was changed at low frequency. Of note, residue 123 and 124 are also important in regulation of vRNA synthesis (Min et al., 2007).

Moreover, NS1 phosphylation sites P215 (Hale et al., 2009), S42 (Hsiang et al., 2012) were 100% conserved, while residue 48 was changed at high frequency in the population. However, as mentioned previously, S48 is not completely conserved among IAVs, and its substitution is thought to have no impact on viral replication (Hsiang et al., 2012).

Of note, contrarily to human H3N2 viruses, no histone mimic sequence was found in EIV NS1 (Marazzi et al., 2012).

Additionally, NS1 PBM changed during evolution (Figure 4-3). Indeed, from emergence to the mid-1990s, NS1 PBM was of the avian type 'ESEV/EPEV' (Obenauer et al., 2006). Then, it then changed to EPEI for a few years before the NS1 C-terminal truncation in the late 1990s. Of note, one virus isolated in 2003 (A/equine/Wisconsin/1/2003) still expressed a full-length version of NS1 with a 'KPKI' motif at residues 227-230. The significance of this motif remains uncertain.

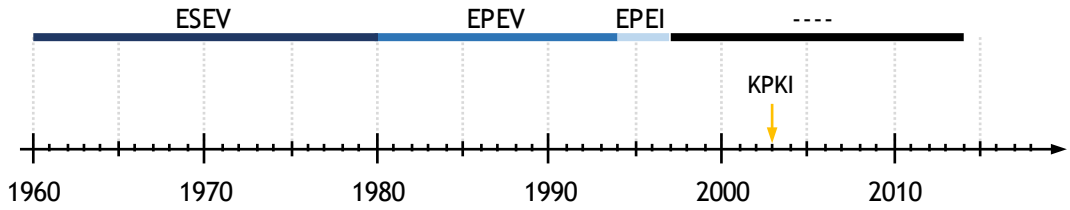


Figure 4-3: Schematic representation of NS1 PBM evolution from 1963 to 2014.

The evolution of NS1 PDZ-binding motif (PBM, residues 227-230) from emergence to 2014 is represented. NS1 PBM was an avian type between 1963 to the early 1990s (ESEV/EPEV) (Obenauer et al., 2006). In the mid-1990s it changed for EPEI, and soon after NS1 C-terminus was truncated of eleven amino acids (represented by '----') resulting in the loss of PBM. One virus isolated in 2003 (A/equine/Wisconsin/2003) still expressed a full-length version of NS1 with a KPKI PBM motif at residue 227-230.

4.2.2. Selection and cloning of phylogenetically distinct NS1 proteins

To experimentally study the functional evolution of EIV NS1 and determine which residue changes identified above could affect protein function, I selected thirteen isolates that included at least one representative virus per decade since 1963 (Table 4-2). I aligned their amino acid sequences and confirmed the presence of the previously described 15 high frequency residue changes (Figure 4-4). Full NS segment, NS1 coding sequence and amino acid sequence for each of the thirteen viruses can be found in Appendices 7 to 19.

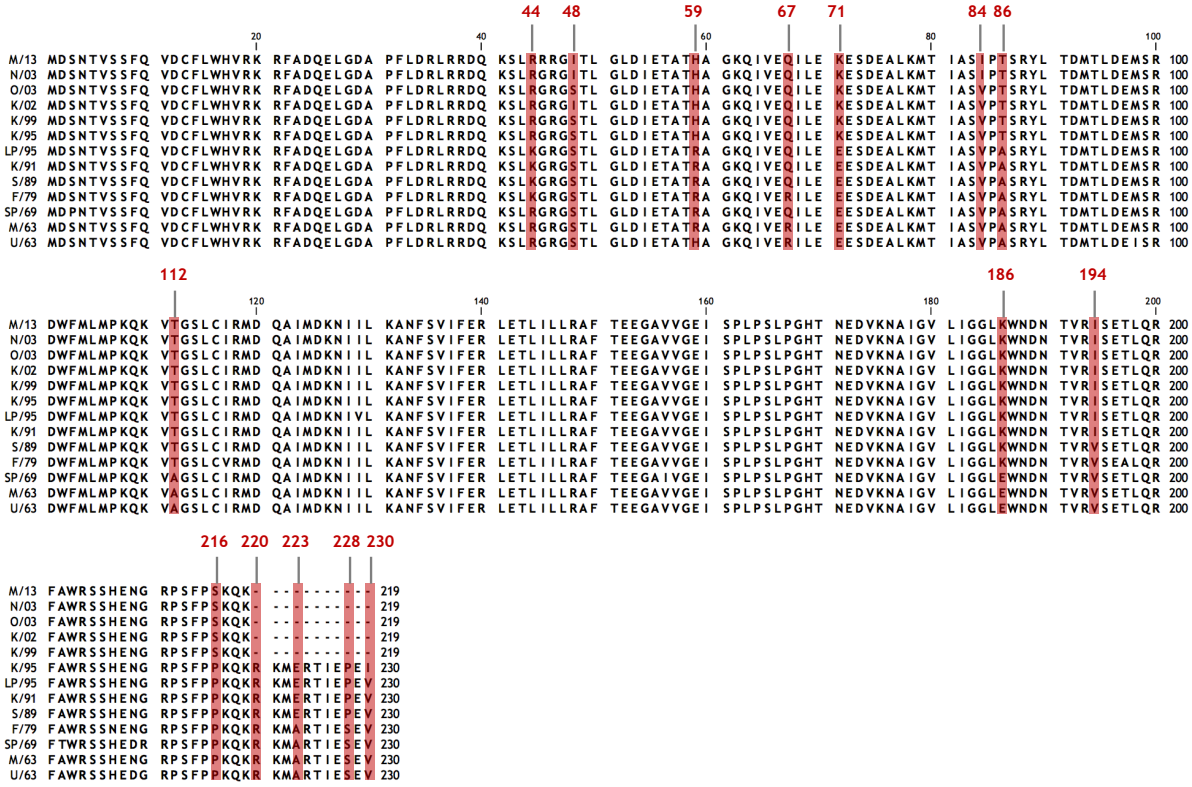
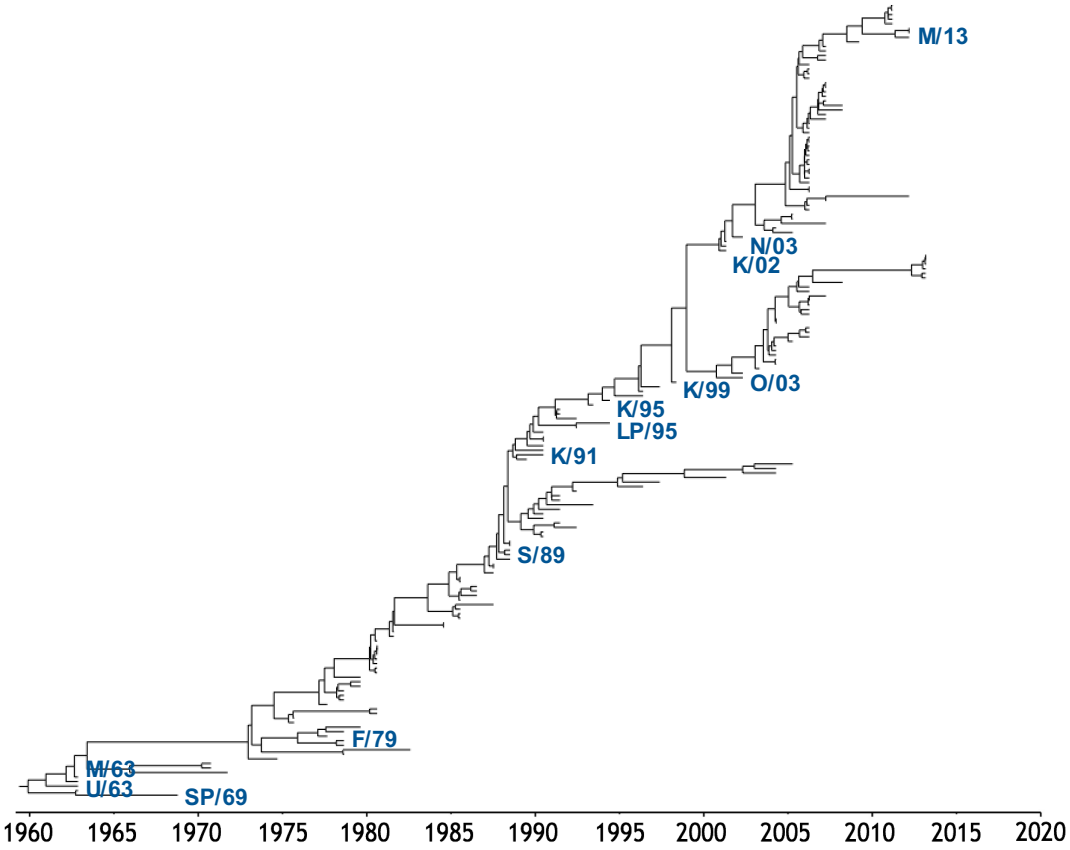


Figure 4-4: Alignment of the NS1 amino acid sequence of EIVs used in this work.

Red shadings highlight high frequency residue polymorphisms in the EIV NS1 population.

Table 4-2: EIVs used in this study.

Virus name	NS segment accession number	Abbreviation
A/equine/Uruguay/1/1963	ACD85423	U/63
A/equine/Miami/1/1963	ABY81497	M/63
A/equine/SaoPaulo/1/1969	ACD85390	SP/69
A/equine/Fontainebleau/1/1979	ACD85401	F/79
A/equine/Sussex/1/1989	ACD97430	S/89
A/equine/Kentucky/1/1991	ACA24672	K/91
A/equine/LaPlata/1995	MF182460	LP/95
A/equine/Kentucky/1995	MF182451	K/95
A/equine/Kentucky/1999	MF182443	K/99
A/equine/Kentucky/5/2002	ABA42429	K/02
A/equine/Ohio/1/2003	ABA42431	O/03
A/equine/Newmarket/5/2003	ACI48802	N/03
A/equine/Mongolia/3/2013	MF182459	M/13



The name, NS segment accession number and abbreviation of the thirteen EIVs used in this work are listed. Their phylogenetic relationship is indicated in blue in the phylogenetic tree obtained as in Figure 4-1.

Each NS1 proteins was then cloned into a pCAGGS expression plasmid through EcoRI and XhoI restriction sites (Figure 4-5-A & -B) and the size of each NS1 insert (693 or 660bp for full length and truncated NS1, respectively) and expression vector (4.7kb) were confirmed by digesting the corresponding plasmid with EcoRI and XhoI restriction enzymes and running the digested product on a 1% Agarose gel (Figure 4-5-C).

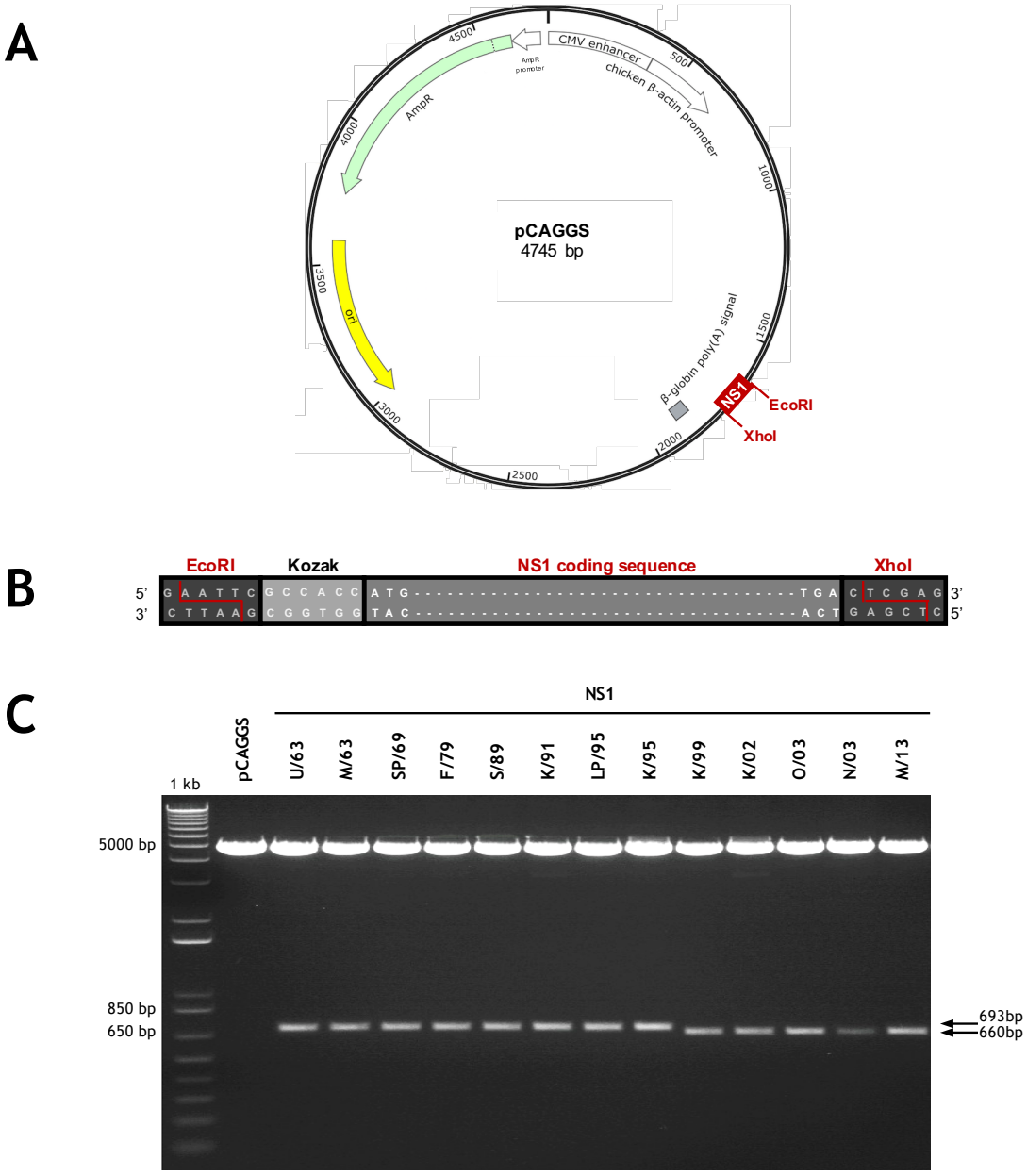


Figure 4-5: Generation of EIV NS1-pCAGGS expression constructs.

A total of thirteen NS1 proteins were cloned into a pCAGGS expression vector through EcoRI and XhoI restriction sites. **(A)** The plasmid map indicates important components of the pCAGGS plasmid as well as the position of the insert. **(B)** A schematic representation the NS1 insert is shown. A Kozak sequence was added after the EcoRI restriction site and the NS1 gene was directly followed by XhoI restriction site. **(C)** The size of each vector (4.7kb) and insert (693 and 660bp, for full length and truncated NS1, respectively) were checked by digesting each plasmid with EcoRI and XhoI, and running the digested products on a 1% agarose gel.

To avoid NEP expression, the splice acceptor site of each NS1 inserts (Figure 4-6-A) was changed from 525-CCAGGA-530 to 525-CCCGGG-530. The size of each NS1 splice-acceptor mutant (SAM) insert and expression vector (Figure 4-6-B) was checked in a similar fashion to Figure 4-4. Finally, the nucleotide sequence of each NS1 insert (WT and SAM) was confirmed by Sanger sequencing (Appendices 4 to 6). A full description of the methodology employed can be found in Chapter 3, section 3.1.

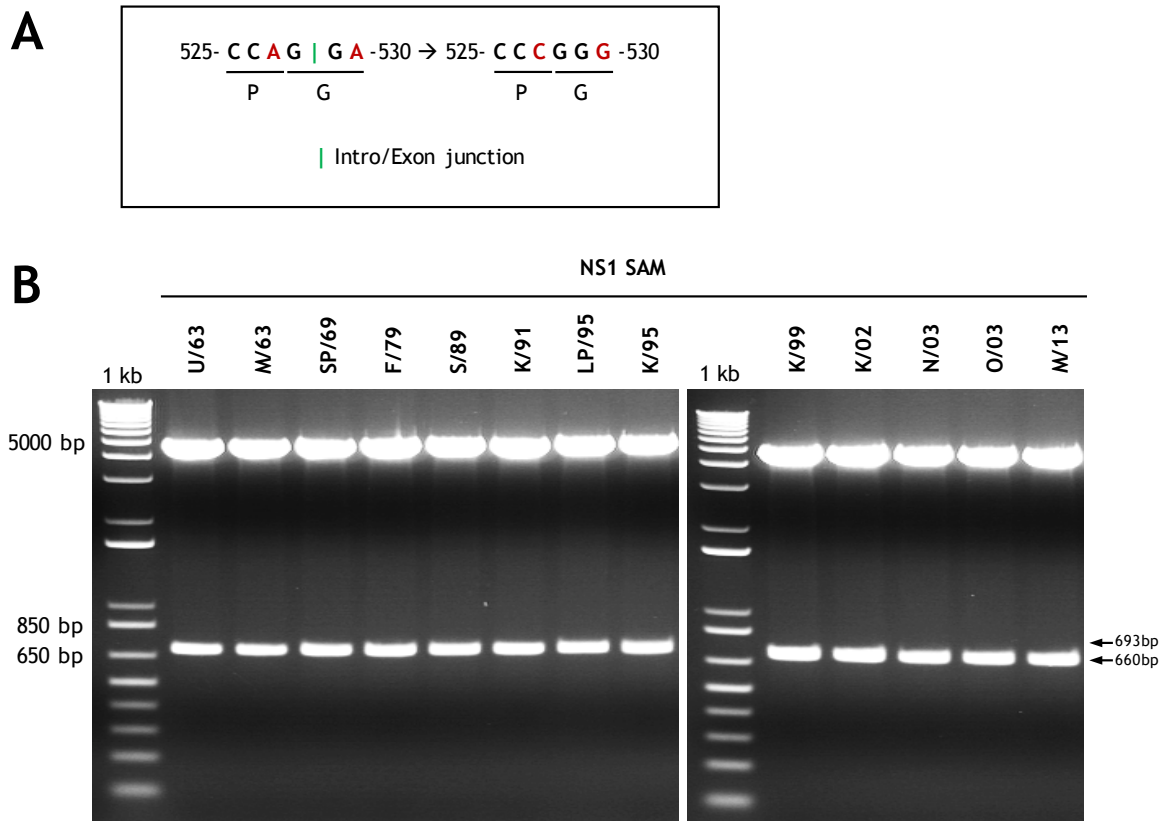


Figure 4-6: Generation of NS1 splice acceptor mutants.

In order to prevent the expression of the nuclear export protein (NEP) in NS1 pCAGGS expression plasmids, a silent mutation in the splice acceptor site (525-CCAG|GA-530) of NS1 was introduced in each insert. **(A)** A schematic representation of the Intro-Exon junction is shown, as well as the amino acid encoded by each codon. **(B)** To check for the size of each NS1 SAM and pCAGGS expression plasmid, NS1 SAM-expressing plasmids were digestion with EcoRI and XhoI restriction enzymes, and the digested products were run on a 1% agarose gel.

4.2.3. Evolution of NS1 function in the EIV lineage

As mentioned previously, NS1 is a well-known blocker of the IFN system via pre-transcriptional and co-/post-transcriptional mechanisms. The gold standard assay to test NS1 function is a reporter assay in which cells are co-transfected with an NS1 expression plasmid together with reporter plasmids expressing a reporter gene (i.e. luciferase gene)

under the control of a promoter of interest (Chapter 3, section 3.2). Then a dual luciferase assay allows the measurement of luciferase activities, which are inversely proportional to the blockade that NS1 exerts on the corresponding promoters. NS1 inhibition of general gene expression (co-/post-transcriptional level) is usually measured by using a reporter plasmid expressing a luciferase gene (i.e. *Renilla* luciferase gene) under the control of a constitutively expressed promoter (pREN-Luc) (Turkington et al., 2015). In addition, to measure NS1's ability to block the production of type I IFN (pre-transcriptional level), the NS1 expressing-vector is co-transfected with a plasmid expressing a reporter gene (i.e. *Firefly* luciferase gene) under the control of the inducible IFN-β promoter (pIFN-β-FF-Luc) (Turkington et al., 2015) together with pREN-Luc, which is then used as an internal control. After a period of incubation, the cells are infected with a virus, i.e. Sendai virus (SeV) (Hale et al., 2010b), to induce the IFN promoter. At the end of the infection period a dual luciferase assay is carried out, and once again, NS1 repression of the inducible IFN-β promoter is inversely proportional to the level of FF-Luc activity detected. Furthermore, the capacity of NS1 to block the induction of ISGs upon IFN treatment can be measured by using a reporter plasmid expressing the *Firefly* luciferase gene under the control of an inducible ISRE-containing promoter (pISRE-FF-Luc). This promoter will be induced by treating the cells with exogenous IFN (i.e. uIFN) (Martinez-Sobrido et al., 2006), and the capacity of NS1 to repress this promoter will again be measured by a dual luciferase assay. Finally, the level of expression of each NS1 protein can be measured by Western blot on parallel transfected cell lysates.

Similar reporter assays were used to study the evolution of EIV NS1 function, and to compare the ability of the thirteen NS1 proteins to block IFN-β and ISRE-inducible promoters, as well as a constitutively expressed promoter. However, prior to this, choosing the most appropriate cell line for this assay was necessary.

4.2.4. Transfection efficiency of two mammalian cell types

The cell type of choice to study EIV NS1 function was equine cells. However, no equine cells from the respiratory tract were available when this work was carried out. Thus, equine dermal fibroblasts (E.Derm/NBL-6, ATCC CCL-57), a commercially available equine cell line previously used to study various equine viruses was chosen (Balasuriya et al., 2014, Soboll Hussey et al., 2011, Maestre et al., 2011). Since their ability to express foreign DNA was unknown and the reporter assay requires cell transfection, their transfection efficiency was

compared to that of HEK293T cells, well known for their high transfection efficiency (Thomas and Smart, 2005), and their faithful protein translation (Ooi et al., 2016). To this end, HEK293T and E.Derm cells were transfected in parallel with a reporter plasmid expressing a green fluorescent protein (GFP) gene under the control of a cytomegalovirus promoter (pCMV-GFP) (Figure 4-7). At 24 hpt the number of cells expressing GFP was analyzed by flow cytometry or fluorescence microscopy. E.Derm cells displayed a very low transfection efficiency, as only 16.3 % [+/-1.0] of cells expressed GFP (Figure 4-7-B). In addition, the examination of the cellular monolayer with a fluorescent microscope confirmed that only a small fraction of cells was expressing GFP (Figure 4-7-A). In contrast, parallel transfections of HEK293T cells confirmed their high transfection efficiency either by fluorescence microscopy (Figure 4-7-C) or by flow cytometry (97.6 % [+/-2.0] of transfected cells expressing GFP) (Figure 4-7-D).

Taken together, these results suggested that HEK293T cells would likely be a better model to test EIV NS1 function in a reporter assay than E.Derm cells.

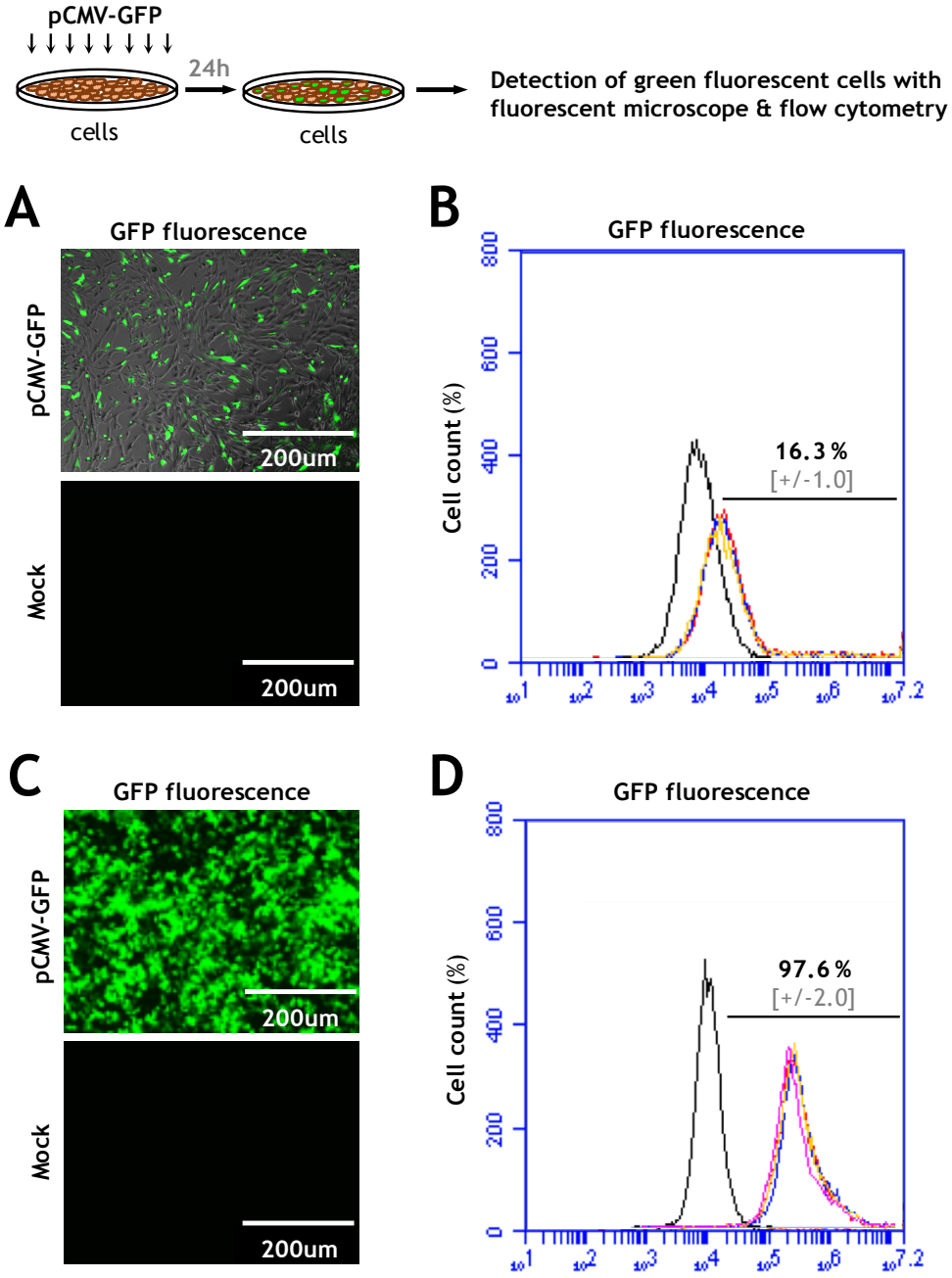


Figure 4-7: Comparison of E.Derm and HEK293T cells transfection efficiency.

(A,B) E.Derm cells and **(C,D)** HEK293T cells were transiently transfected with a pCMV-GFP expression plasmid (1000ng) for 24h as described in the method section. At the end of the transfection period the cellular monolayer was **(A,C)** examined with a fluorescent microscope to confirm the presence of green fluorescent cells, and A representative picture of each condition is shown. **(B,D)** Cells were then detached and fixed in Formalin as indicated in the method section, and the number of green fluorescent cells was determined by flow cytometry. The experiment was repeated three times independently and means and error margins are indicated by the black and grey values respectively.

4.2.5. Function of phylogenetically distinct NS1 proteins

First, the ability of the thirteen NS1 proteins to limit the production of IFN-β was compared. To this end, HEK293T cells were co-transfected with three different amounts of NS1 expression plasmids (1000ng, 200ng, 20ng) or with an empty vector (pCAGGS, 1000ng), together with 50ng of pIFN-β-FF-Luc, and with 50ng of pREN-Luc used for normalization (Figure 4-8). Total plasmid concentration was kept constant (1000ng) with an empty expression vector. At 24 hpt, the cells were infected with 50 HAU of SeV for a further 18 h in order to stimulate the IFN-β promoter, and at the end of the infection period, both luciferase activities were measured. As shown in (Figure 4-8-A), SeV infection resulted in high increase of IFN-β-FF-Luc activity in cells co-transfected with the empty vector (set to 100%) compared to uninfected controls. However, the IFN-β promoter was strongly antagonized in a dose-dependent manner in cells expressing each NS1 proteins tested.

The ability of these NS1 proteins to block induction of ISGs upon IFN treatment was then compared by co-transfecting HEK293T cells with NS1 expressing plasmids (1000ng, 200ng and 20ng), together with 50ng of the pISRE-FF-Luc reporter plasmid, and pREN-Luc. At 24 hpt, cells were treated with 500U of uIFN to stimulate the ISRE-containing promoter, and 18 h later ISRE-FF-Luc and REN-Luc activities were measured. Interestingly, the NS1 proteins of EIVs isolated in 1963 (U/63 and M/63) displayed a relatively low repression of the ISRE-containing promoter (Figure 4-8-B). NS1's blockade of this promoter progressively increased between 1969 and 1995 (SP/69 to K/95), while most NS1 from 1999 to 2013 (K/99 to M/13) showed a dose-dependent control of this promoter.

To compare the ability of the thirteen NS1 protein to inhibit general gene expression, individual NS1 expression plasmids (1000ng, 200ng, 20ng) were co-transfected together with pREN-Luc (50ng), and REN-Luc activity was measured 24 h later. *Renilla* luciferase activity was strongly antagonized by NS1 proteins from the 1960s (U/63, M/63, SP69) (Figure 4-8-C). In contrast, NS1 from 1979 onwards (F/79 to M/13) were unable to block general gene expression, as shown by high REN-Luc expression levels. Furthermore, for some of these recent NS1 proteins an increase in gene expression was observed (increase in REN-luc activity) (Figure 4-8-C). These results were consistent with the western blot results (Figure 4-8-D), as NS1 proteins that strongly antagonized the constitutively expressed promoter (i.e. U/63, M/63, SP/69) could not be detected, while the ones that did not (F/79 to M/13) were easily detectable. Interestingly, the level of expression of NS1 of viruses isolated between 1979 and 1995 (F/79 to LP/95) was lower than that of more recent NS1 proteins (K/95 to M/13).

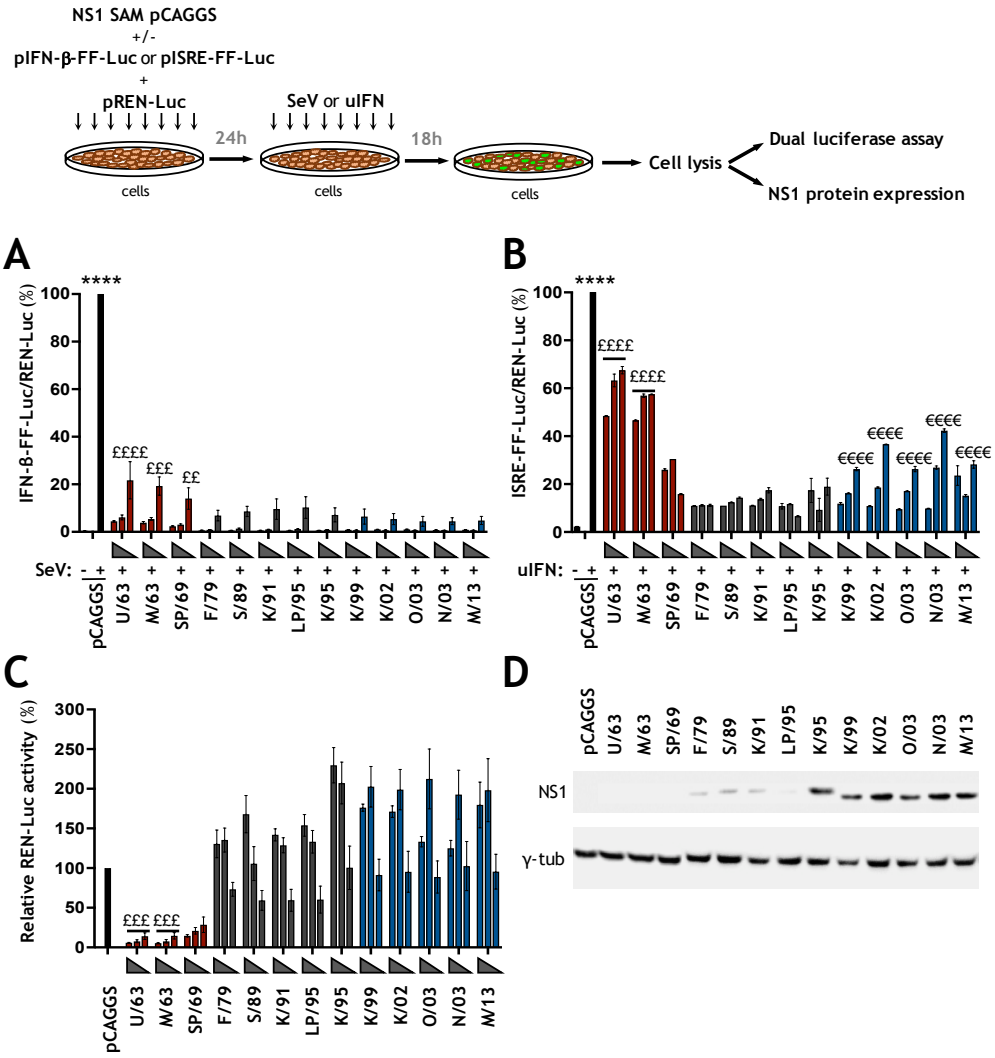


Figure 4-8: Functional characterization of evolutionary distinct EIV NS1 proteins in HEK293T cells.

HEK293T cells were transiently co-transfected with a pCAGGS expression plasmid (1000, 200, 20ng) encoding the indicated NS1 proteins (or empty plasmid) together with either **(A)** a *Firefly* luciferase (FF-Luc)-IFN-b promoter reporter construct (pIFN-b-FF-Luc), or **(B)** a FF-Luc-IFN-stimulated regulatory element reporter construct (pISRE-FF-Luc), as well as a constitutively active *Renilla* luciferase expression plasmid (pREN-Luc). At 24 hpt, cells were either **(A)** infected with SeV (+) or mock infected (-), or **(B)** treated with universal interferon (uIFN) (+) or mock treated (-) for 18h. **(A-B)** The relative activity of FF-Luc was determined as the ratio between FF-Luc and REN-Luc in each corresponding sample. Values were normalized to **(A)** empty pCAGGS plasmid (+) SeV or **(B)** empty pCAGGS plasmid (+) uIFN (set to 100%). **(C)** HEK293T cells were transiently co-transfected individually with the indicated NS1-expression plasmids (1000, 200, 20ng) together with pREN-Luc. Total REN-Luc levels were measured 24 h later, and values were normalized to empty pCAGGS transfected cells. **(A, B and C)** Bars correspond to mean of three independent experiments and error bars represent SEM. A Two-way ANOVA with Bonferonni post hoc test was used to determined significance. **(A and B)** ****, $p<0.0001$ for pCAGGS+SeV or +uIFN (1000, 200, 20ng) against all other conditions; **(A)** £££££, $p<0.0001$ for U/63 versus F/79 to M/13; £££, $p<0.001$ for M/13 versus F/79 to M/13; ££, $p<0.01$ for SP/69 versus F/79 to M/13 (20ng); **(B)** £££££, $p<0.0001$ for U/63 and M/63 versus F/79 to M/13 (1000, 200, 20ng); €€€€, $p<0.0001$, K/99 to M/13 versus U/63 to K/95 (20ng); **(C)** £££, $p<0.001$ for U/63 and M/13 versus F/79 to M/13 (1000, 200ng). **(D)** The level of expression of the indicated NS1 protein was analysed by Western blot on parallel transfected lysates (1000ng) using a rabbit polyclonal anti-EIV NS1 antibody. The Western blot was run 3 times independently and a representative picture is shown.

Despite their low transfection efficiency, similar reporter assay were tested in E.Derm cells. Three NS1 expression plasmids (U/63, K/95, O/03) were selected based on their phenotype in HEK293T cells (Figure 4-8). E.Derms were then transfected with 1000ng of plasmid together with 500ng of pIFN-β-FF-Luc or pISRE-FF-Luc, and 500ng of pREN-Luc used for normalization (Figure 4-9-A & 4-9-B). At 24 hpt, cells were either infected with 50 HAU of SeV or treated with 500U of uIFN, to induce the IFN-β promoter or the ISRE-containing promoter, respectively. At 18 h post infection/treatment, a dual luciferase assay was I carried out. In addition, to assess NS1 capacity to block general gene expression, the three NS1 expression plasmids (1000ng) were co-transfected with pREN-Luc (500ng), and REN-Luc activity was measured 24 h later (Figure 4-9-C). Finally, the level of expression of the three NS1 proteins was analyzed by western blot from parallel transfected cell lysates (Figure 4-9-D).

As shown in (Figure 4-9-A & -B), similar results to what was found in HEK293T cells were observed in E.Derm, where all NS1 proteins blocked efficiently IFN-β induction (Figure 4-9-A). However, since the level of induction of the IFN-β promoter was very low, these results are controversial. In addition, as obtained with HEK293T cells, only U/63 NS1 efficiently blocked the constitutively expressed promoter (Figure 4-9-C). However, contrarily to what was found in HEK293T cells, only K/95 could efficiently repress the ISRE-containing promoter in E.Derm (Figure 4-9-B). Moreover, none of the NS1 protein tested could be detected by Western blot (Figure 4-9-D), which was likely due to the low transfection efficiency of these cells.

Hence, HEK293T cells were confirmed to be a more reliable model for this experiment.

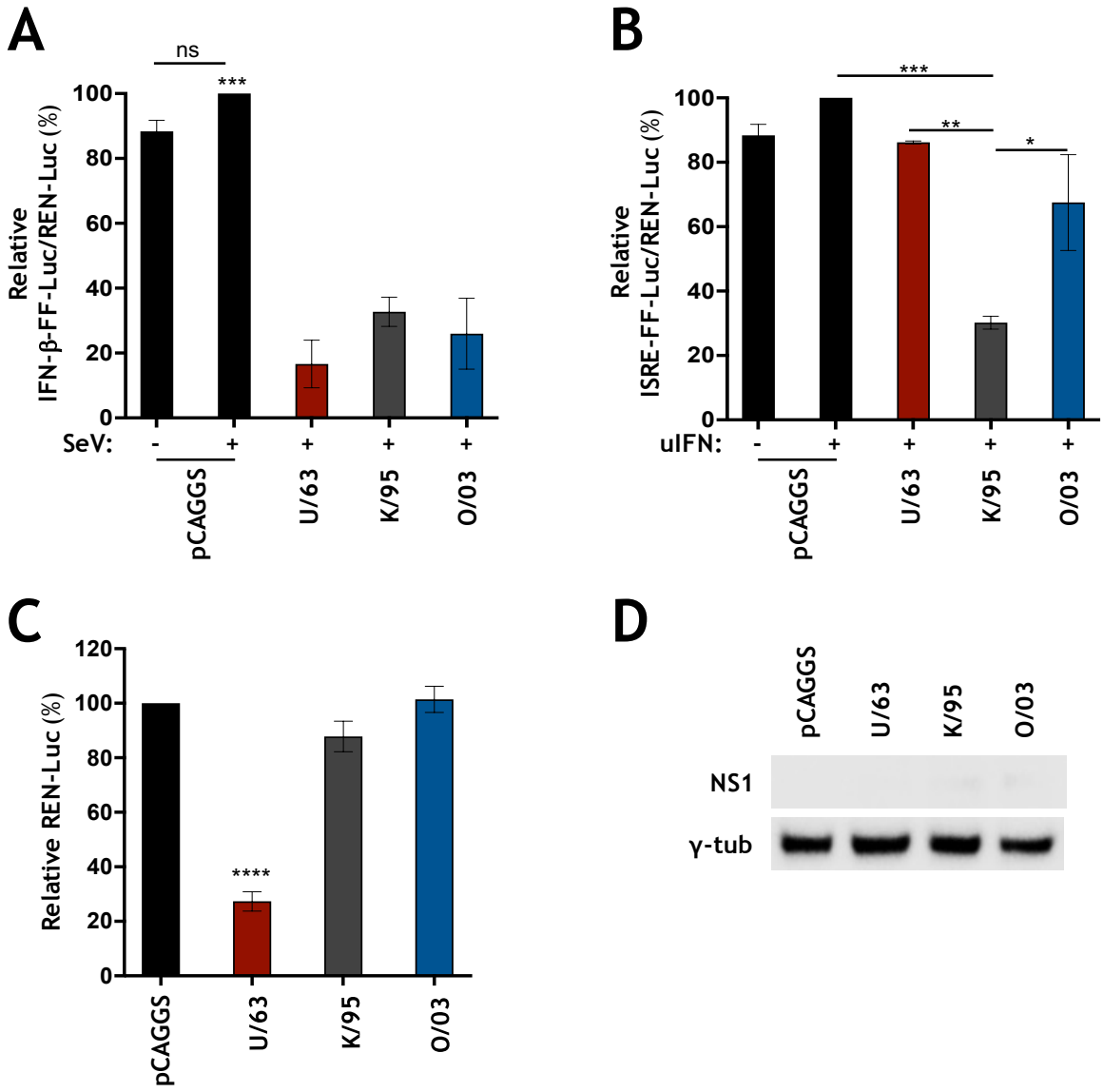


Figure 4-9: Functional characterization of evolutionary distinct EIV NS1 proteins in E.Derm cells.

E.Derm cells were transiently co-transfected with a pCAGGS expression plasmid (1000ng) encoding the indicated NS1 proteins (or empty plasmid), together with either **(A)** pIFN-b-FF-Luc (500ng), or **(B)** pISRE-FF-Luc (500ng), as well as pREN-Luc (500ng). At 24 hpt, cells were either **(A)** infected with SeV (+) or mock infected (-), or **(B)** treated with universal interferon (uIFN) (+) or mock treated (-) for 18h. **(A-B)** The relative FF-Luc activities were determined as for Figure 4-7. **(C)** E.Derm cells were transiently co-transfected with the indicated NS1-expression plasmids (1000ng) together with pREN-Luc. Total REN-Luc levels were determined as for Figure 4-7. **(A-C)** Bars correspond to mean of three independent experiments and error bars represent SEM. Significance was calculated as for Figure 4-7: **(A)** ***, $p<0.001$ for pCAGGS-SeV versus U/63 to O/03; **(B)** ***, $p<0.001$ for K/95 versus pCAGGS+SeV; **, $p<0.01$ for K/95 versus U/63; *, $p<0.05$ for K/95 versus O/03. **(C)** ****, $p<0.0001$ for U/63 versus all other conditions. **(D)** The level of expression of the indicated NS1 was analysed as for Figure 4-7.

4.2.6. Impact of residues 112 and 186 on NS1 function

As seen in (Figure 4-8-C) and (Figure 4-9-C), the NS1 protein of EIVs that circulated in the 1960s (U/63, M/63, SP69) strongly blocked the otherwise constitutively expressed *Renilla* luciferase, while the NS1 protein of EIVs isolated after 1979 (F/79 to M/13) did not. Interestingly, in the early 1970s two amino acid substitutions (E186K and A112T) arose in EIV NS1 (Figure 4-2) Since residue 186 lies within the NS1 CPSF30-binding domain, and by binding to CPSF30 NS1 can block host gene expression (Ayllon and Garcia-Sastre, 2015, Hale et al., 2010a), E186K substitution was strongly suspected to be responsible for the observed change of function.

However, to thoroughly test which substitution(s) was(were) responsible for the observed change, residue 186 and 112 single and double NS1 mutants were generated for two phylogenetically distinct EIVs, which naturally harbor A112 and E186 (U/63), and T112 and K186 (O/03) (Figure 4-10). The three mutant versions of each protein were generated by mutagenesis (Figure 4-11 & 4-12), and the presence of the desired mutation(s) was checked by Sanger sequencing (Figure 4-11-A & 4-11-B & Figure 4-12-A & 4-12-B). In addition, the size of each insert and vector were checked prior to use (Figures 4-11-C & 4-12-C).

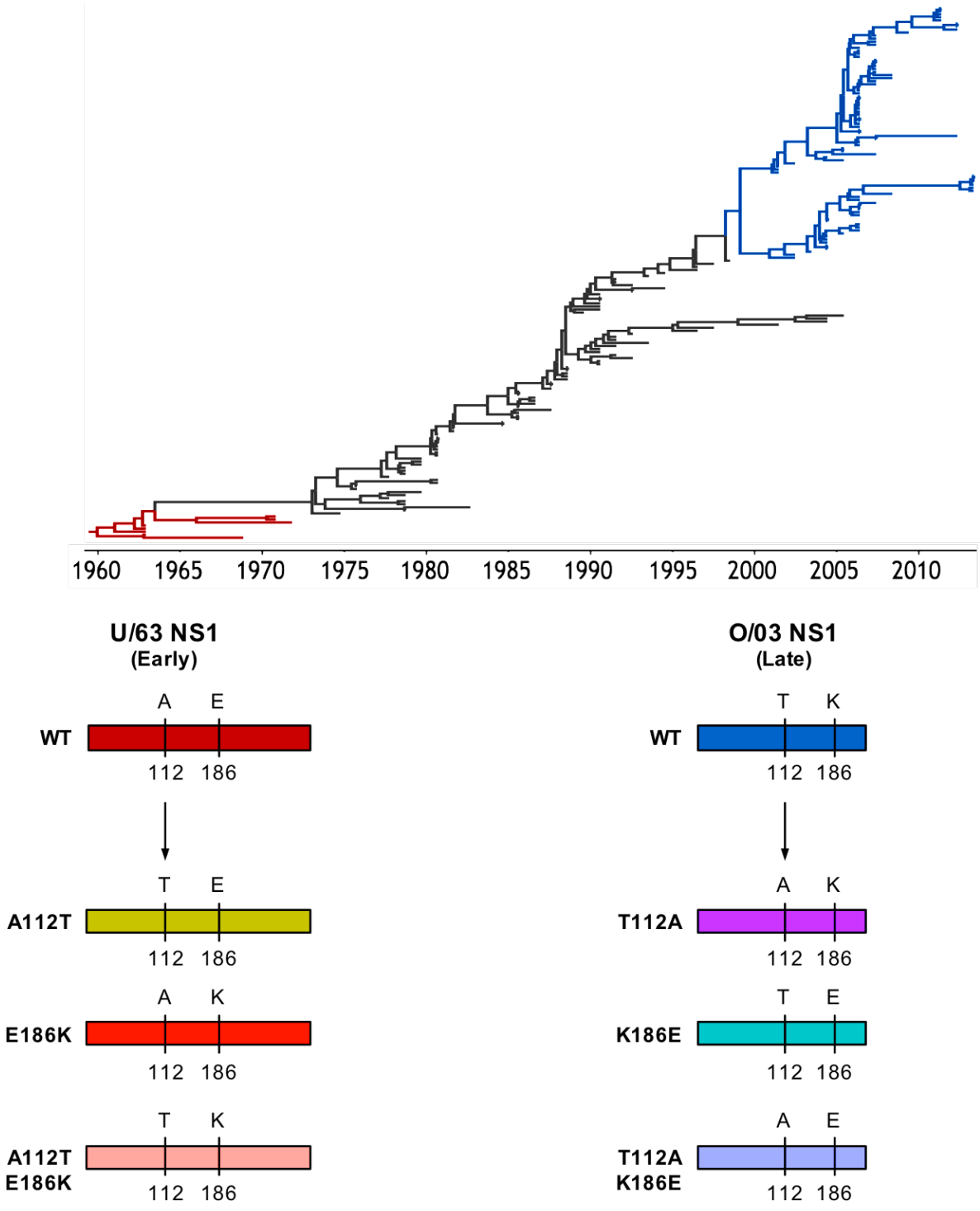


Figure 4-10: Generation of residue 112 & 186 NS1 mutants.

Two phylogenetically distinct EIV NS1 proteins (U/63 and O/03 NS1), which naturally contain either A or T at position 112 and E or K at position 186 were selected. Single and double mutants of these NS1 proteins were generated. A schematic representation of each NS1 WT and mutant is shown below the tree.

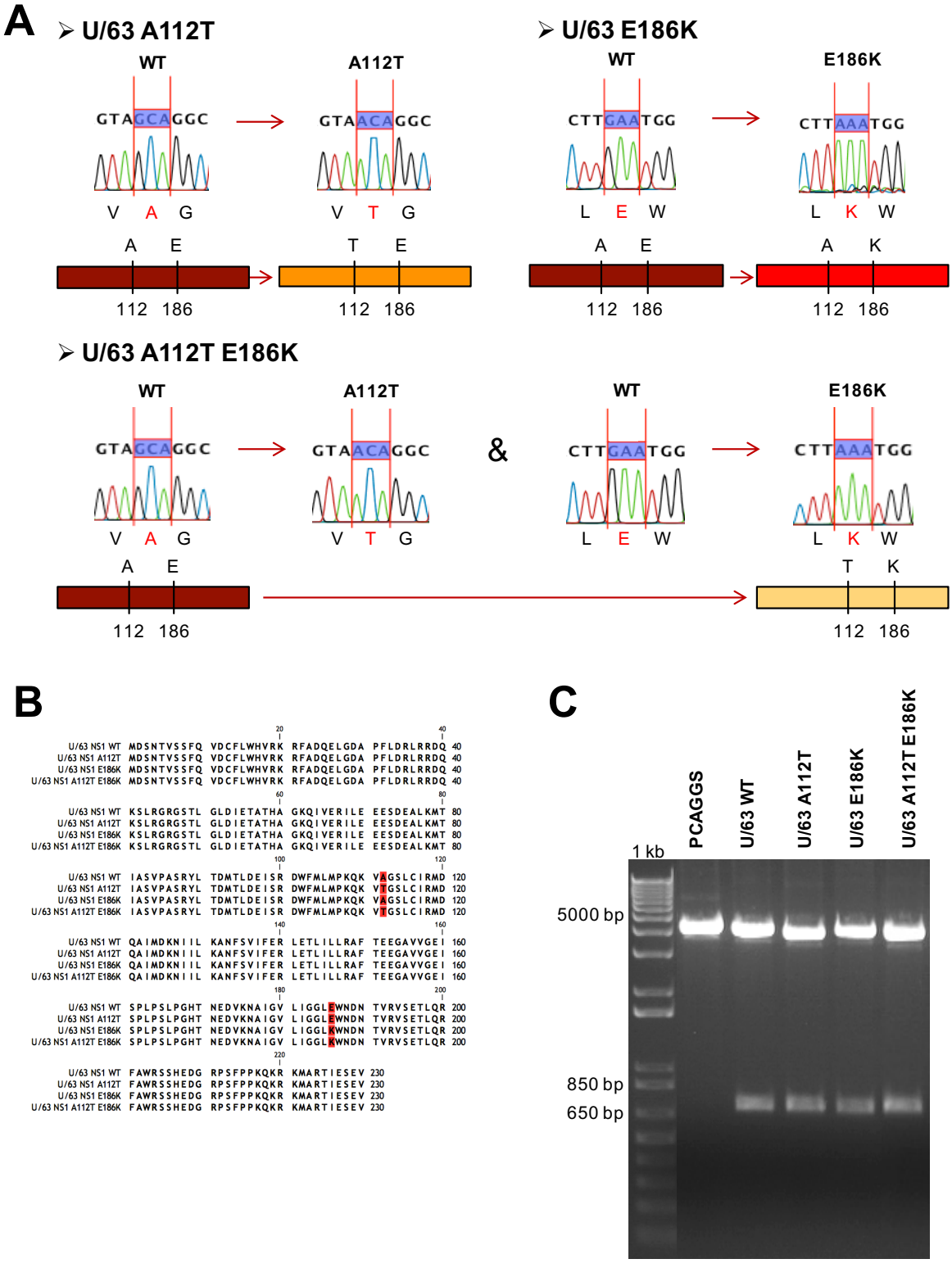


Figure 4-11: Generation of U/63 NS1 mutants for residue 112 and 186.

(A) NS1 WT, single and double mutant inserts were sequenced by Sanger sequencing. Corresponding traces are shown. **(B)** Alignment of the amino acid sequence of WT and mutant NS1 inserts are shown. Red shadings indicate a different amino acid. **(C)** Each NS1 (WT and mutant) expression plasmid was digested with EcoRI and XhoI restriction enzymes and digested products were ran on a 1% agarose gel to verify plasmid (4.7kb) and insert (693bp) sizes.

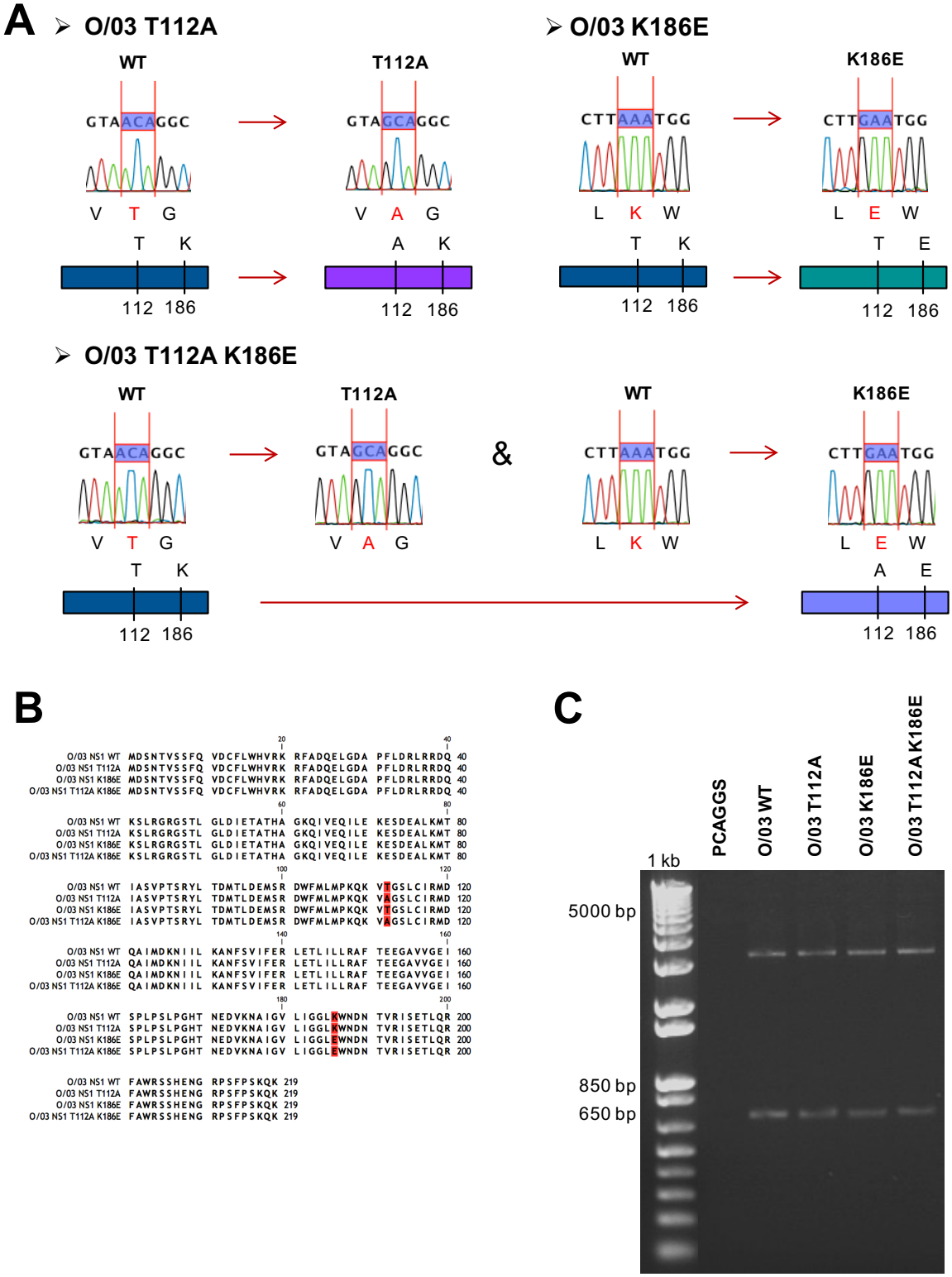


Figure 4-12: Generation of O/03 NS1 mutants for residue 112 and 186.

As for Figure 4-10, **(A)** NS1 WT, single and double mutant inserts were sequenced by Sanger sequencing. Corresponding traces are shown. **(B)** Alignment of the amino acid sequence of WT and mutant NS1 inserts are shown. Red shadings indicate a different amino acid. **(C)** Each NS1 (WT and mutant) expression plasmid was digested with EcoRI and XhoI restriction enzymes and digested products were ran on a 1% agarose gel to verify plasmid (4.7kb) and insert (660bp) sizes.

As expected, none of the introduced mutations affected NS1 control of IFN-β induction (Figure 4-13-A), while NS1 proteins harboring E186 (i.e. U/63 WT, U/63 A112, O/03 K186E and O/03 T112A K186E) strongly repressed general gene expression, and K186-containing NS1 proteins (i.e. U/63 E186K, U/63 EA112T E186K, O/03 WT and O/03 T112A) did not (Figure 4-13-C). Consistent with this the NS1 proteins harboring E186 were not detectable by western blot. In addition, the expression of K186-containing U/63 NS1 protein remained lower than that of K186-containing O/03 NS1 proteins (Figure 4-13-D). Surprisingly, the presence of K186 in U/63 NS1 proteins did not fully rescue the phenotype observed for K186-containing O/03 NS1 proteins. This suggested that U/63 NS1 control of general gene expression is not fully due to E186, and other mutations are likely involved in this function.

Additionally, the presence of K186 seemed to be associated with a better control of the ISRE-containing promoter, as K186-containing NS1 proteins (i.e. U/63 E186K, U/63 A112T E186K and O/03 WT or O/03 T112A) repressed more pISRE-FF-Luc activity than their E186-containing counterparts (i.e. U/63 WT, U/63 A112, O/03 K186E and O/03 T112A K186E) (Figure 4-13-B).

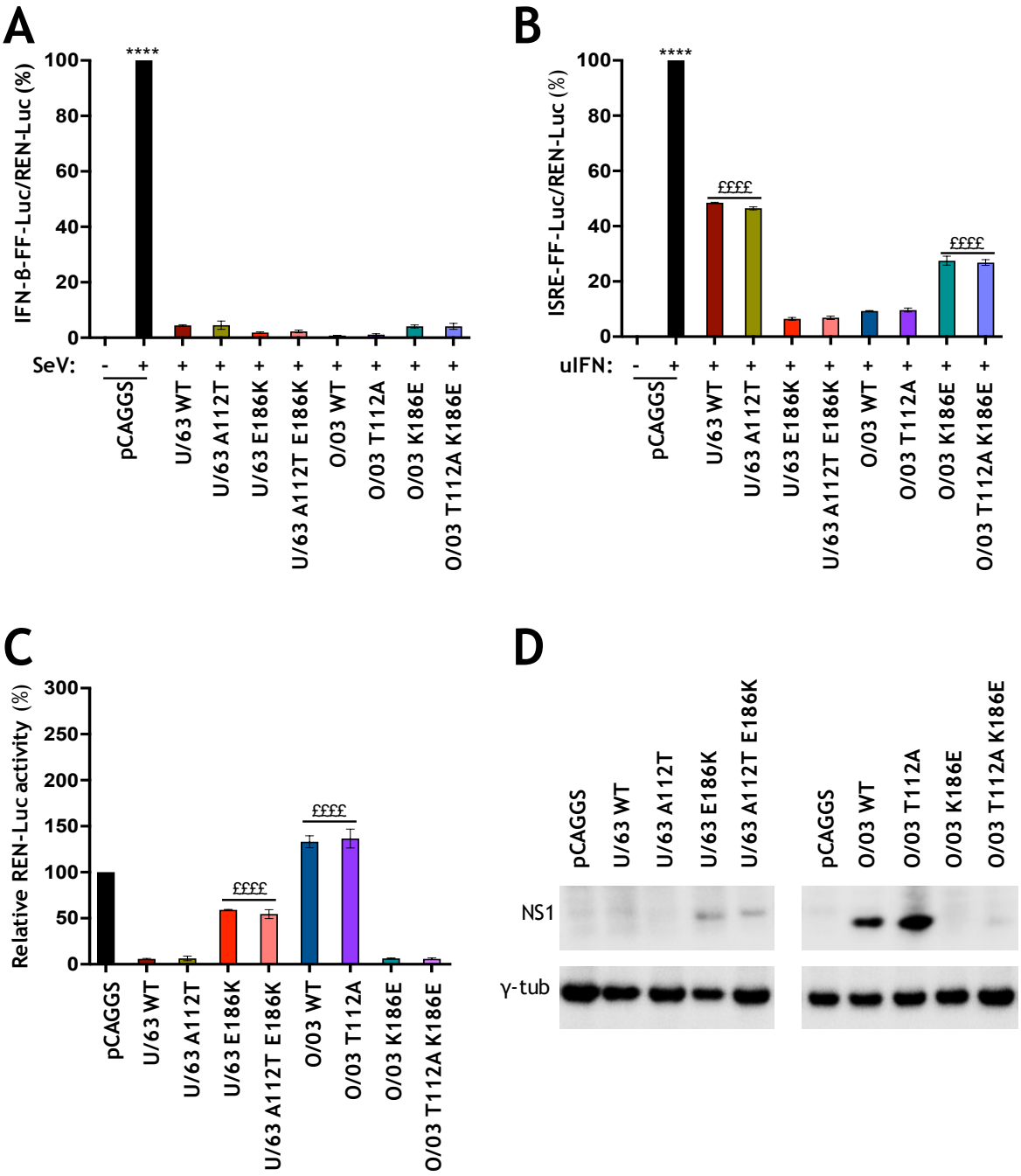


Figure 4-13: Functional characterization of U/63 and O/03 NS1 112 & 186 mutants in HEK293T cells.

As for Figure 4-7, HEK293T cells were transiently co-transfected with a pCAGGS expression plasmid (1000ng) encoding the indicated NS1 proteins (or empty plasmid), together with either **(A)** pIFN-b-FF-Luc (50ng), or **(B)** pISRE-FF-Luc (50ng), as well as with pREN-Luc (50ng). **(A-B)** Inducible promoters were activated, and the relative FF-Luc activities were determined as for Figure 4-7. **(C)** Human 293T cells were transiently co-transfected with the indicated NS1-expression plasmids (1000ng) together with pREN-Luc (50ng), and total REN-Luc levels were determined as for Figure 4-7. **(A-C)** Bars correspond to mean of three independent experiments and error bars represent SEM. Significance was calculated as for Figure 4-7: **(A and B)** ****, $p<0.0001$ for pCAGGS+SeV or +uIFN versus all other conditions; ££££, $p<0.0001$ for indicated NS1 contructs versus all others. **(D)** The level of expression of the indicated NS1 was analyzed by Western blotting as for Figure 4-7.

4.2.7. Impact of residue 186 and C-terminus on NS1 function

Another factor than residue 186 seemed to affect NS1 control of gene expression and ISG induction, as O/03 NS1 phenotype could not be fully rescued by introducing K186 in U/63 NS1 (Figure 4-13-B, 4-13-C & 4-13-D). Interestingly, another major difference between U/63 and O/03 NS1 was their respective length. Indeed, as other EIVs isolated between 1963 and the early 1970s, U/63 expressed a full length NS1 protein, which is composed of 230 amino acids. In contrast, as for EIVs isolated after the late 1990s, O/03 contained a short version of NS1, truncated of 11 amino acids at its C-terminus. (Figure 4-14). As a consequence, U/63 NS1 possesses an 'ESEV' PBM motif and a PABPII binding site, while O/03 NS1 does not. These motifs could allow U/63 NS1 to control signal transduction (Obenauer et al., 2006) and general gene expression (Li et al., 2001, Chen et al., 1999) either independently or in complement to E186. To test this hypothesis, three phylogenetically distinct EIVs displaying functional differences in the reporter assay (Figure 4-8) were selected. U/63, K/95, and O/03, possessing the early (E186, 230 amino acids), intermediate (K186 230 amino acids) and late (K186, 219 amino acids) version of NS1, respectively (Figure 4-14), were used to generate single and double mutants for residue 186 and C-terminus (Figure 4-15, 4-16, & 4-17).

As for earlier constructs, the presence of the desired mutation(s) was verified by Sanger sequencing (Panel A & B of Figure 4-15, 4-16, & 4-17), and the size of each insert and vector were checked on a 1% agarose gel (Panel C of Figure 4-15, 4-16, & 4-17) prior to the reporter assay.

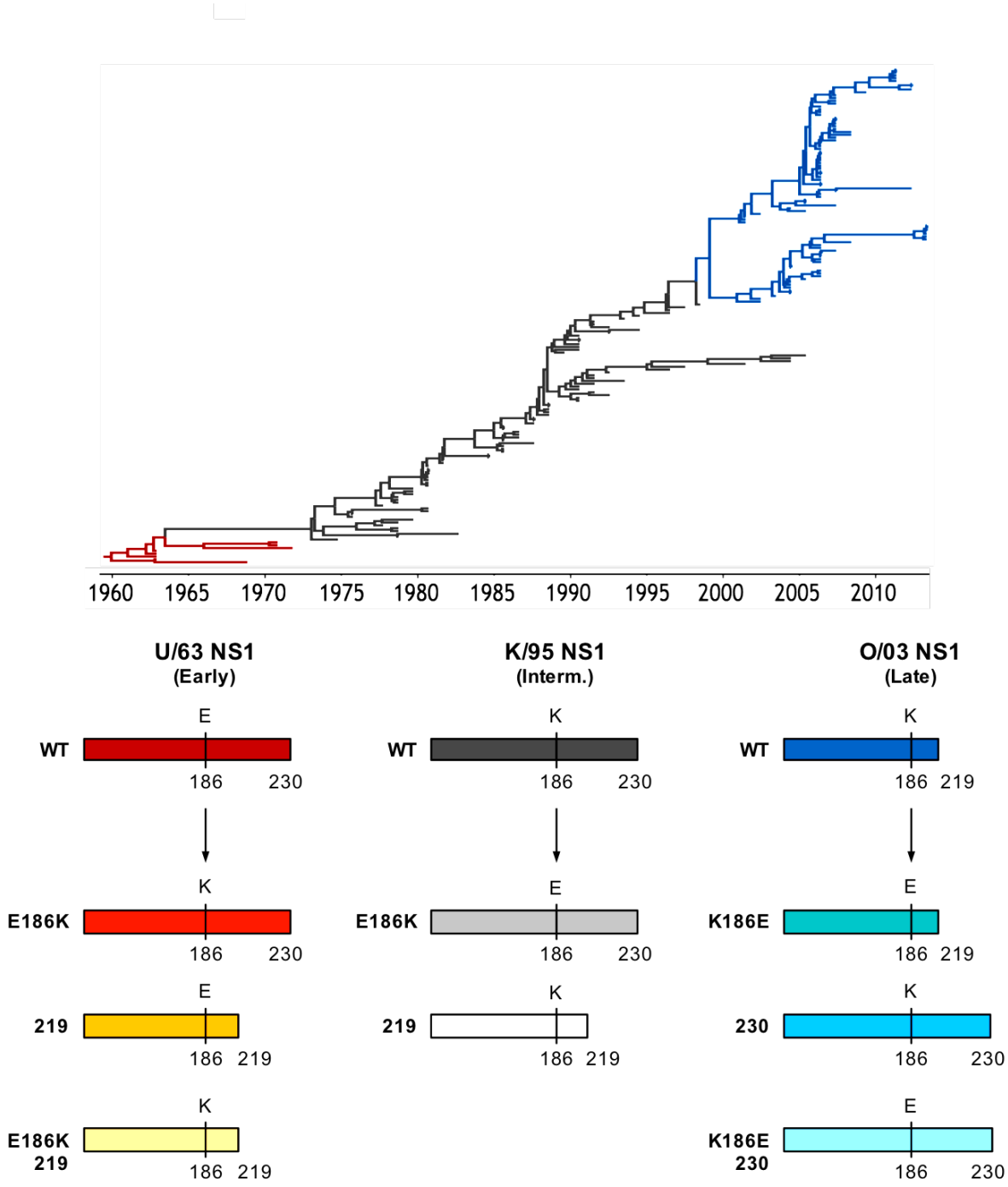


Figure 4-14: Generation of residue 186 & C-terminus NS1 mutants.

Three phylogenetically distinct EIVs (U/63, K/95 and O/03), which naturally contain either E or K at position 186 and long and short C-terminal tail (230 or 219 amino acids, respectively) were selected, and single or double mutant versions of these NS1 proteins were generated. A schematic representation of NS1 WT and mutants is shown.

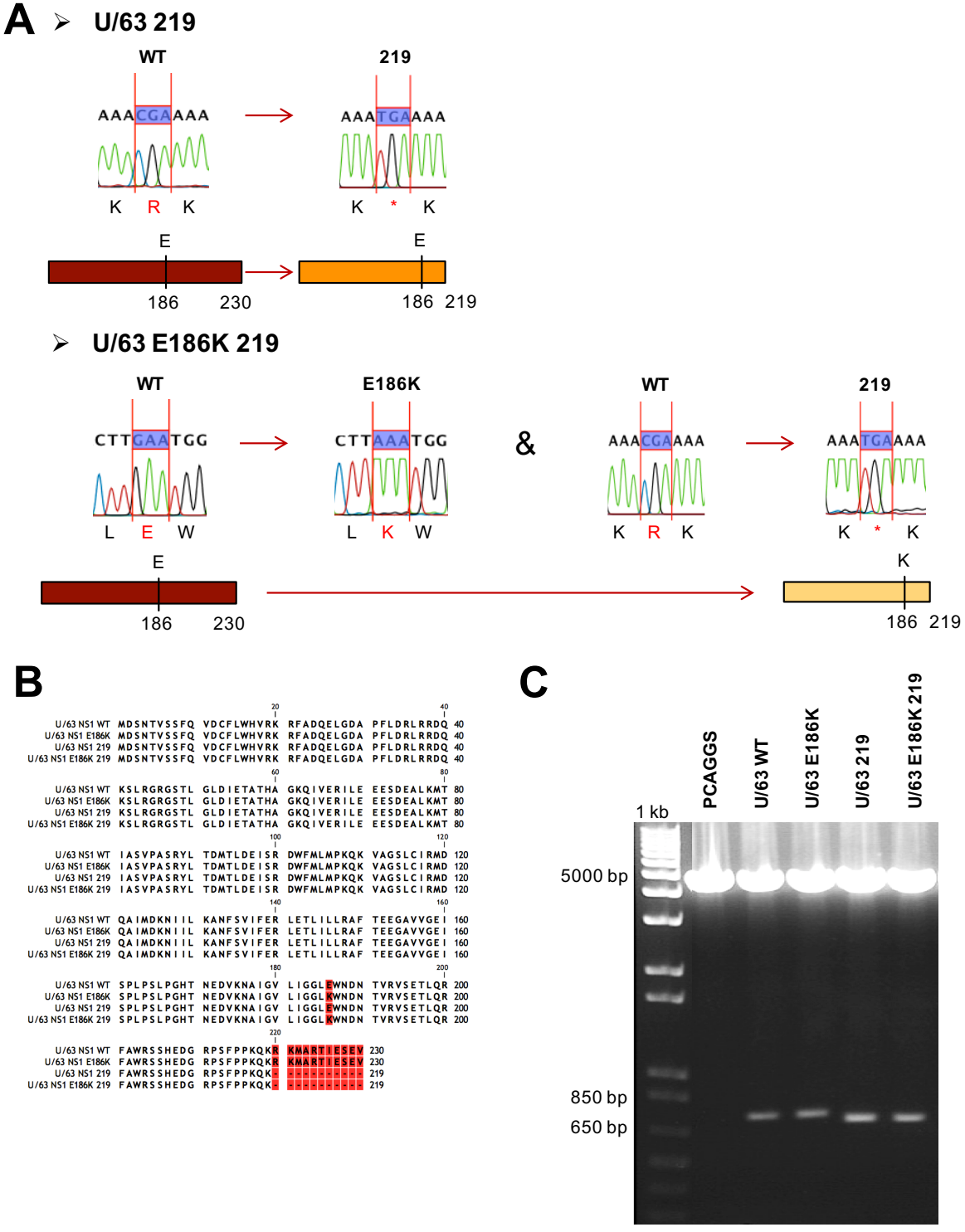


Figure 4-15: Generation of U/63 NS1 mutants for residue 186 and C-terminus.

As for Figure 4-10, **(A)** NS1 WT, single and double mutant inserts were sequenced by Sanger sequencing. Corresponding traces are shown. **(B)** Alignment of the amino acid sequence of WT and mutant NS1 inserts are shown. Red shadings indicate a different amino acid. **(C)** Each NS1 (WT and mutant) expression plasmid was digested with EcoRI and XhoI restriction enzymes and ran on a 1% agarose gel to verify plasmid (4.7kb) and insert (660 and 693bp) sizes.

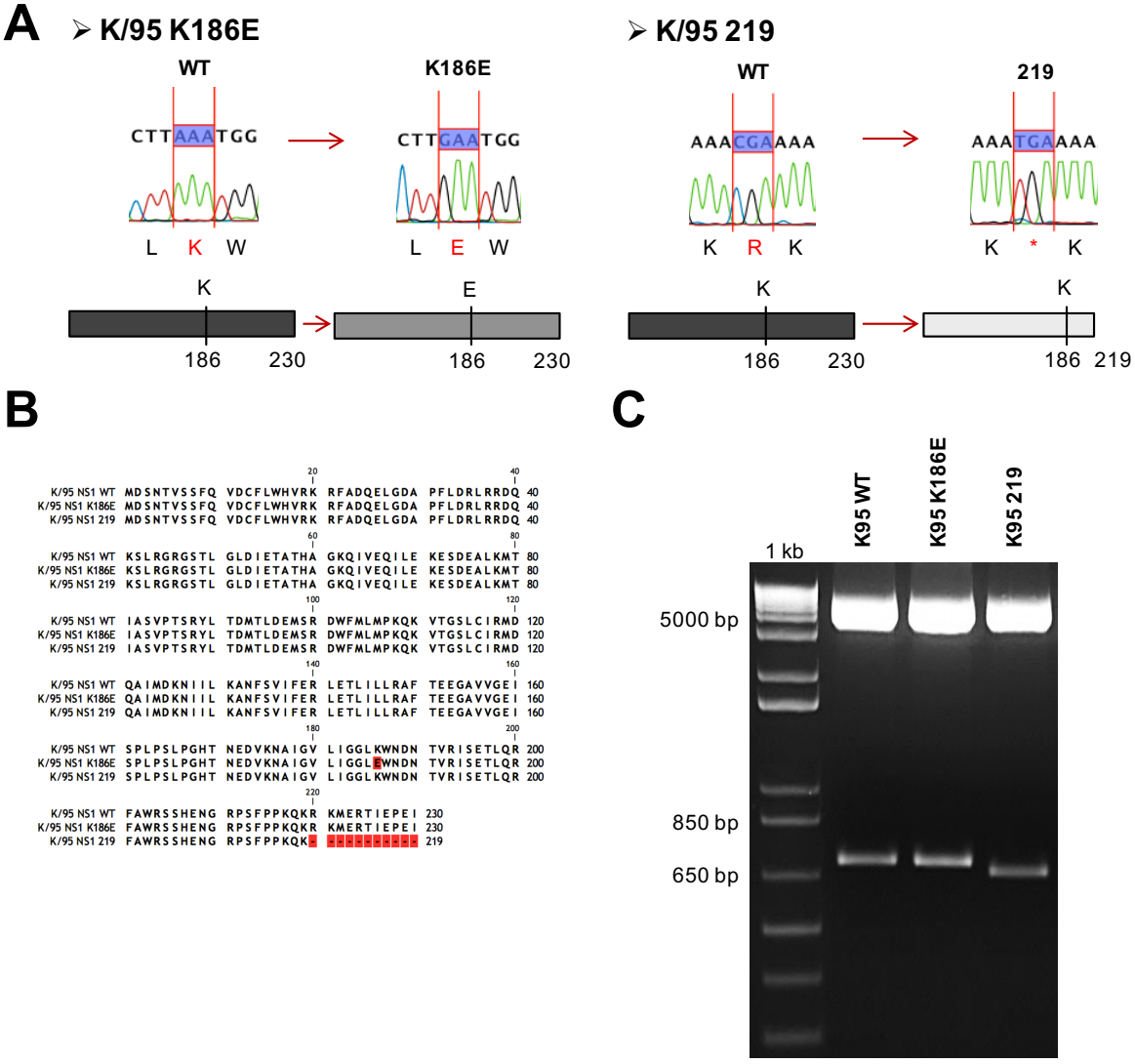


Figure 4-16: Generation of K/95 NS1 mutants for residue 186 and C-terminus.

As for Figure 4-10, **(A)** NS1 WT, single and double mutant inserts were sequenced by Sanger sequencing. Corresponding traces are shown. **(B)** Alignment of the amino acid sequence of WT and mutant NS1 inserts are shown. Red shadings indicate a different amino acid. **(C)** Each NS1 (WT and mutant) expression plasmid was digested with EcoRI and XhoI restriction enzymes and ran on a 1% agarose gel to verify plasmid (4.7kb) and insert (660 and 693bp) sizes.

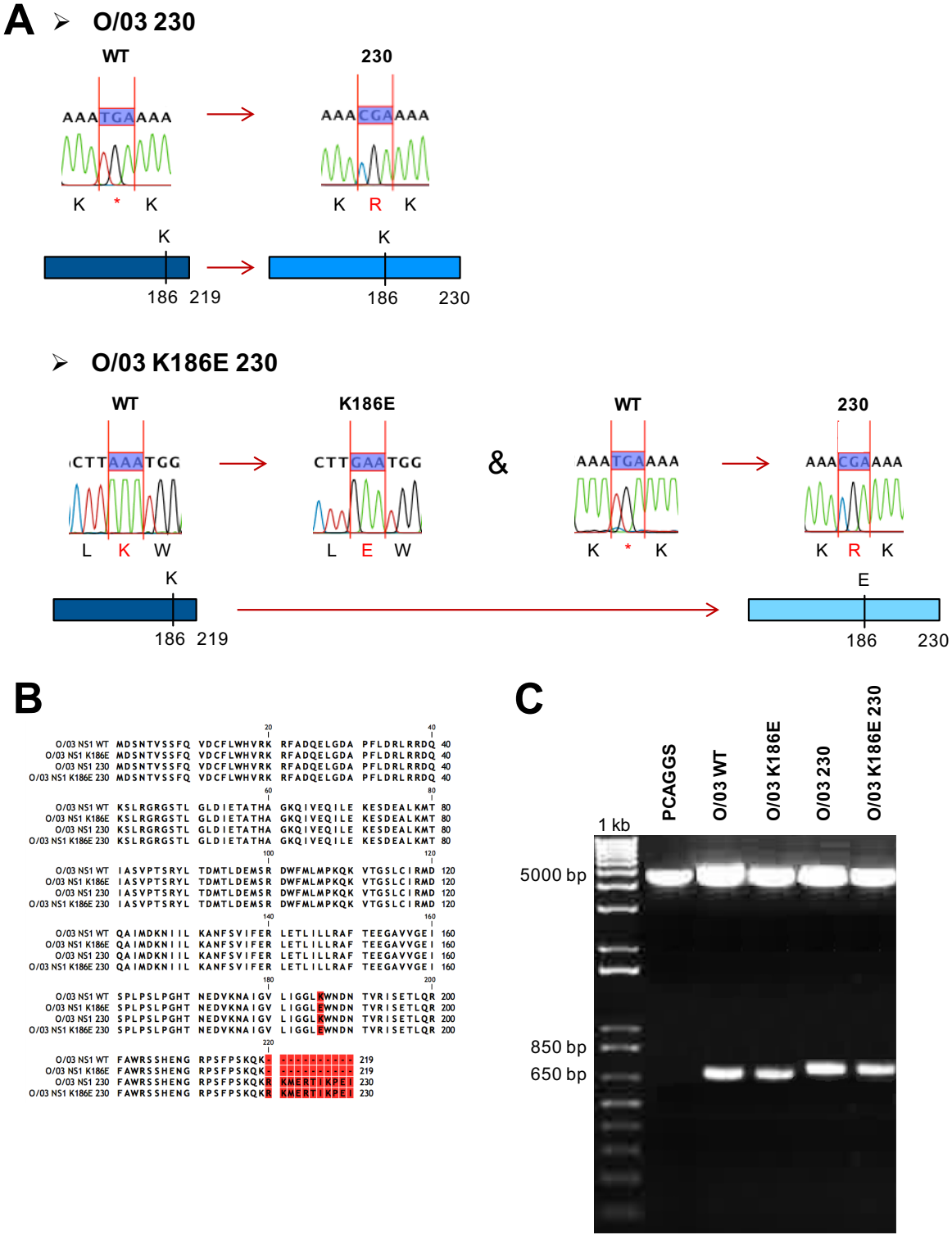


Figure 4-17: Generation of O/03 NS1 mutants for residue 186 and C-terminus.

As for Figure 4-10, **(A)** NS1 WT, single and double mutant inserts were sequenced by Sanger sequencing. Corresponding traces are shown. **(B)** Alignment of the amino acid sequence of WT and mutant NS1 inserts are shown. Red shadings indicate a different amino acid. **(C)** Each NS1 (WT and mutant) expression plasmid was digested with EcoRI and XhoI restriction enzymes and ran on a 1% agarose gel to verify plasmid (4.7kb) and insert (660 and 693bp) sizes.

Their ability to block IFN-β and ISG induction, and to control general gene expression was then assessed in a reporter assay (Figure 4-18). Consistent with previous results, none of the introduced mutations affected NS1 control of IFN-β induction (Figure 4-18-A). In contrast, the presence of E186 in NS1 was associated with a strong repression of general gene expression (Figure 4-18-C), and these proteins were not detectable by western blot (Figure 4-18-D). Also, K186-containing-K/95 and -O/03 NS1 proteins strongly induced REN-Luc expression (Figure 4-18-C) and were strongly expressed in transfected cells (Figure 4-18-D). Furthermore, residue 186 seemed to affect NS1 control of ISG induction (Figure 4-18-B), as all E186-containing NS1 proteins were controlling less efficiently the ISRE-containing promoter compared to their K186 counterparts. These results were consistent with observation made in Figure 4-13. Surprisingly, K/95-K186E mutant lost the ability to control the ISRE-containing promoter when transfected at 200ng (middle dose) (Figure 4-18-B). Whether this result is artefactual or real would require further investigation.

The effect of NS1 C-terminus on protein function seemed to be context-dependent (Figure 4-18-B). Indeed, in the context of U/63, the C-terminal truncation (U/63-219) slightly increased NS1 control over the ISRE-containing promoter, particularly at middle and low doses (200ng, 20ng). However, when introduced in K/95 (K/95-219), no effect was observable at low dose (20ng) and a reduction of control was observable at middle and high doses (1000ng, 200ng) compared to WT (Figure 4-18-B). Finally, for O/03 NS1 no difference was observable for the ISRE-containing promoter, but it seemed to stimulate general gene expression at high dose (1000ng) (Figure 4-18-C).

Of note, no synergistic effect between 186 substitution and C-terminal truncation/extension was observable in this assay.

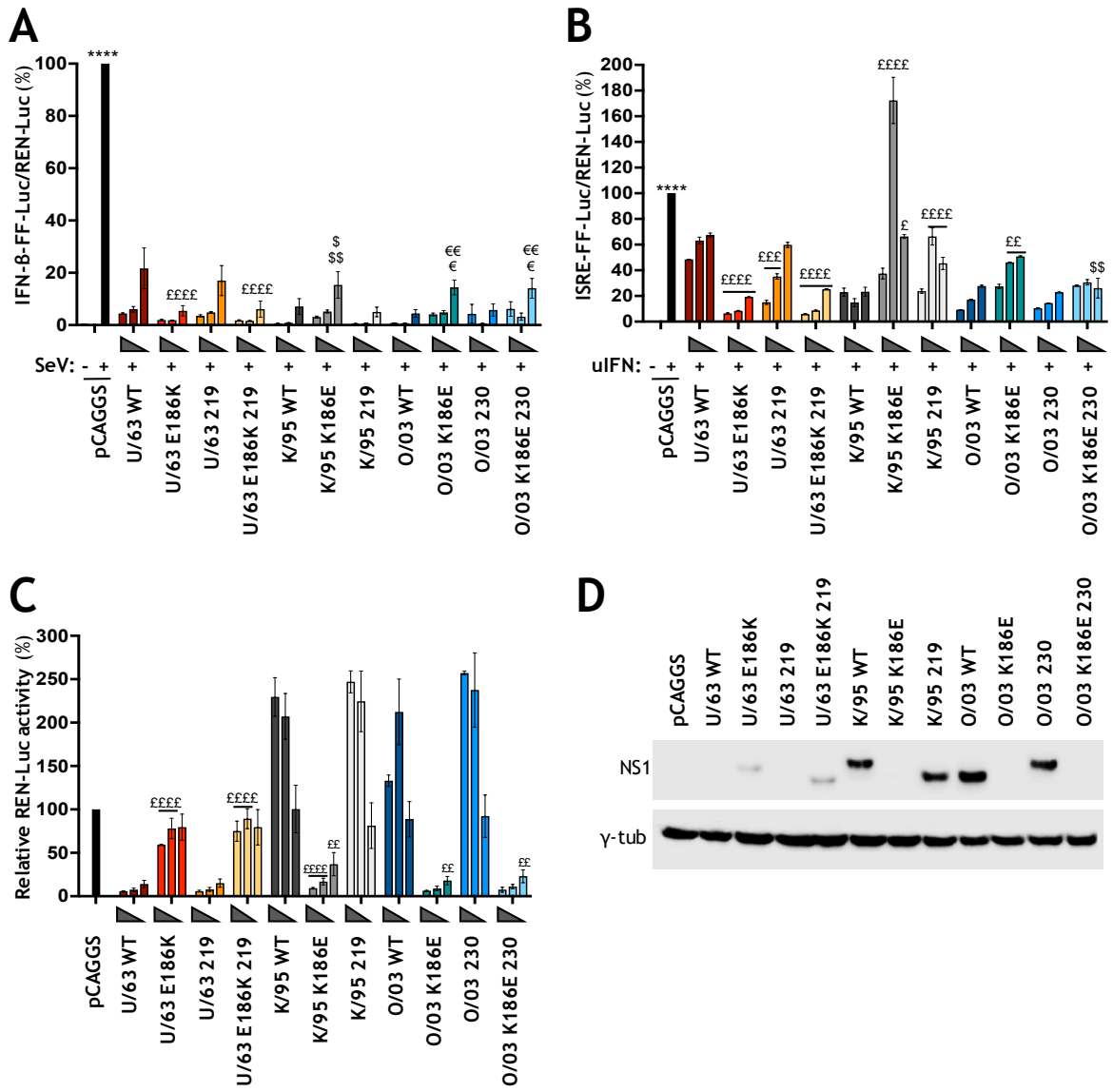


Figure 4-18: Functional characterization of U/63, K/95 and O/03 NS1 186 and C-terminus mutants in HEK293T cells.

As for Figure 4-7, HEK293T cells were transiently co-transfected with a pCAGGS expression plasmid (1000ng, 200ng, 20ng) encoding the indicated NS1 proteins (or empty plasmid), together with either **(A)** pIFN-b-FF-Luc (50ng), or **(B)** pISRE-FF-Luc (50ng), as well as with pREN-Luc (50ng). **(A-B)** Inducible promoters were activated, and the relative FF-Luc activities were determined as for Figure 4-7. **(C)** HEK293T cells were transiently co-transfected with the indicated NS1-expression plasmids (1000ng, 200ng, 20ng) together with pREN-Luc (50ng), and total REN-Luc levels were determined as for Figure 4-7. **(A-C)** Bars correspond to mean of three independent experiments and error bars represent SEM, and significance was calculated as for Figure 4-7: **(A and B)** ****, p<0.0001 for pCAGGS+SeV or +uIFN versus all NS1 constructs and pCAGGS-SeV or –uIFN. **(A)** ££££, p<0.0001 for U/63-WT & -219 versus U/63-E186K-containing NS1s; $, p< 0.05 for K/95-K186E versus K/95 WT & $$, p< 0.01 for K/95-K186E versus K/95-219; **(B)** £££ to £, p<0.0001 to p<0.05 for indicated dose of NS1 construct against similar dose of respective NS1 WT construct. $$, p<0.01 for O/03-K186E-230 against O/03-K186E. **(D)** NS1 level of expression was analysed by Western blotting.

4.2.8. Impact of residue 186 and C-terminus on NS1-CPSF30 interaction

Since, residue 186 is located in the middle of NS1-CPSF30 binding interface (185-GLEWND-189), and this binding has previously been associated with the blockade of general gene expression (Ayllon and Garcia-Sastre, 2015, Hale et al., 2010a), it became evident to test if residue 186 was influencing NS1-CPSF30 interaction. To do so, co-immunoprecipitation assays between an equine CPSF30 and the O/03 NS1 WT and mutants were performed (Figure 4-19). As suspected, the presence of E186 was associated with O/03 NS1 interaction with CPSF30, independently of the protein length.

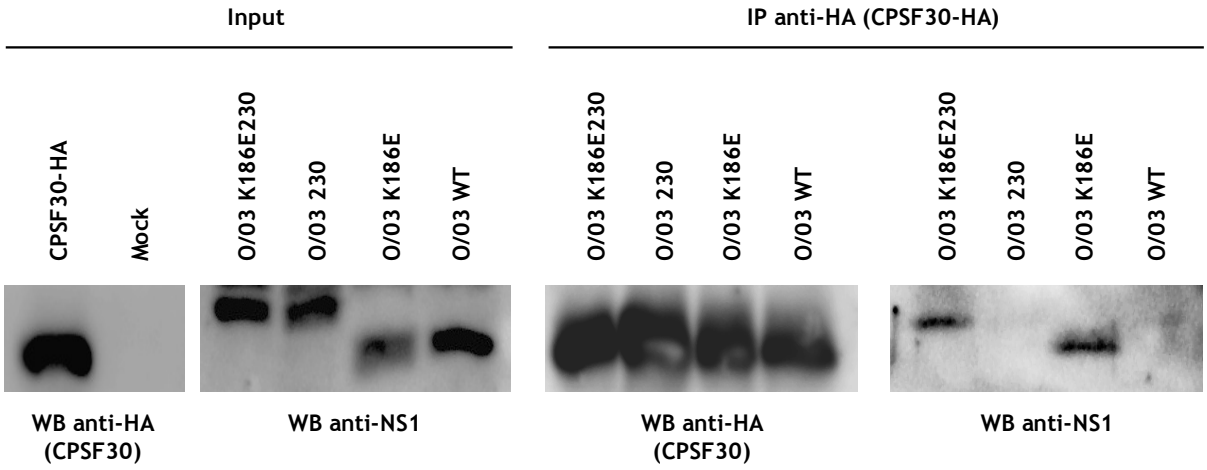


Figure 4-19: Evaluation of residue 186 as a determinant of NS1-CPSF30 interaction.

The 4 O/03 NS1 variants were synthesized *in vitro*, and HEK293T cells were transiently transfected with 2000ng of a pCAGGS plasmid expressing a HA-tagged version of the human CPSF30. At 30 hpt, cells were lysed and cleared cell lysates expressing HA-CPSF30 were incubated overnight with the *in vitro*-synthesized NS1 proteins and 20 µl of an anti-HA affinity resin, as described in the method section. Precipitated proteins were dissociated from the resin and analysed by Western blotting. The experiment was repeated three times independently and representative pictures of protein expression levels ('input', left panel) and co-immunoprecipitation results (IP anti-HA, right panel) are shown.

4.3. Discussion

In this chapter, the evolution of H3N8 EIV NS1 during fifty years of circulation of the virus in the equine population was studied. By analyzing the phylogenetic relationship of 175 NS1 genes of the EIV lineage (Figure 4-1), and by comparing the amino acid sequence of these proteins (Appendices 1 & 2), 15 amino acid changes fixed throughout evolution were identified: K44R, S48I, R59H, R67Q, E71K, V84I, A86T/V, A112T, E186K, V194I, P216S, R220*, A223E, S228P, V230I (Table 4-1). These findings correlated with previous reports from the literature identifying residue 67 (Murcia et al., 2011b) and substitution S228P and V194I (Murcia et al., 2011a) as point of selection pressure and host marker sites in the EIV genome, respectively. In addition, substitutions 44, 59, 71, 86, 216 and C-terminal truncation were previously identified as genetic signatures of viruses of the Florida clades currently circulating in the equine population (Barba and Daly, 2016).

Interestingly, apart from substitution K44R that reverted to the original amino acid (R44K), all changes were fixed at the population level (Table 4-1 & Figure 4-2). More importantly, the clades that did not introduce them died out shortly after separation from the main trunk (Figure 4-1 & 4-2). Although further work would be needed to determine their exact contribution to NS1 function, this suggests that these residue changes may play a role in EIV fitness and must have been selectively advantageous for the virus. In addition, other residue changes may have occurred at the population level and could have remained below the threshold of detection due to an insufficient number of virus isolates.

Additionally, several mutations seemed to have occurred synchronously in the EIV population (i.e. A112T and E186K; R44K reversion and R59H, E71K, A86T, V230I substitutions; P216S substitution and R220*) (Figure 4-2). However, their role in NS1 function and their impact on NS1 evolution remain uncertain. Their apparition could have been the result of the so-called 'genetic hitchhiking' (Morley and Turner, 2017), wherein a beneficial mutation sweeps to fixation and carries along any other neutral or mildly deleterious mutations occurring on the same genetic background. This is likely the case for E186K and A112T substitution, as residue 112 did not seem to affect NS1 function (Figure 4-13). Alternatively, two independent beneficial mutations may have occurred on the same genetic background and sweep to fixation together. Furthermore, if two or more of these mutations were beneficial in combination it could have driven genomes containing them to rise in frequency. It would be interesting to evaluate in which of these categories the high-frequency NS1 residue changes belong.

To move forward with the study of EIV NS1 function evolution, the ability of 13 phylogenetically distinct NS1 proteins to block general gene expression, and to control type I IFN and ISG induction were tested using a reporter assay (Figure 4-8). It was observed that during the first 10 years post-emergence, EIV NS1 strongly repressed general gene expression (Figure 4-8-C), due to the presence of a glutamic acid at position 186 (Figure 4-13-C & 4-18-C), which in turn allowed NS1 to interact with CPSF30 (Figure 4-19). These early NS1 proteins also displayed a relatively low repression on ISG induction (Figure 4-8-B). The introduction of E186K substitution in the early 1970s led to a loss of shutoff of general gene expression (Figure 4-13-C & 4-18-C) due to a loss of CPSF30-binding (Figure 4-19). These findings are in accordance with previous reports from the literature showing that specific substitutions destabilizing the NS1-CPSF30 complex were reported to arise when viruses adapted to replicate in certain host species, i.e. duck to quail (Hossain et al., 2008), or human to mouse (Brown et al., 2001). Also, other groups also reported several nonhuman-adapted virus strains (either through laboratory passages or natural selection) encoding NS1 proteins that lacked the ability to interact with CPSF30 (Hale et al., 2010b), i.e. the egg- and mouse-adapted PR8 (Hayman et al., 2006, Kochs et al., 2007b), the highly pathogenic H5N1 and its presumed precursor H6N1 (Twu et al., 2007), and the swine-origin 2009pH1N1 (Hale et al., 2010b).

The E186K substitution was also associated with an increased repression of ISG induction (Figure 4-8-B, 4-13-B). Interestingly, residue 186 is directly neighboring W187, known to regulate NS1 ED dimerization (Hale et al., 2008, Aramini et al., 2011, Kerry et al., 2011), and to strengthen dsRNA binding, thus reinforcing NS1 anti-IFN activity (Aramini et al., 2011). This ED dimerization was also shown to be incompatible with the CPSF30 binding (Kerry et al., 2011). Thus, it is possible that the loss of CPSF30 binding ten years post emergence has allowed EIV NS1 to interact with new binding partners or unmasked or reinforced important functions that were selectively advantageous for the virus.

However, the normalization used in these reporter assays was suboptimal. Indeed, the promoter of a control vector should be compatible with the experimental conditions and provide consistent expression under the experimental conditions being tested. To test NS1's repression of IFN-β and ISRE promoters pREN-Luc was used for normalization (Figure 4-8-A&B, 4-9-A&B, 4-13-A&B, and 4-18-A&B), but variability in NS1's control over pREN-Luc was notably observed between E186- versus K186-containing NS1 proteins (Figure 4-8-C, 4-9-C, 4-13-C, and 4-18-C). Furthermore, no internal control was used when testing NS1's control over general gene expression (Figure 4-8-C, 4-9-C, 4-13-C, and 4-18-C).

For our transfection experiments, it would have been preferable to use an eukaryotic cell system expressing the T7 RNA polymerase and a T7 promoter control (Nogales et al.,

2017). Indeed, this expression system is independent of the RNA pol II and is not controlled by NS1.

Interestingly, the control over IFN-β induction was strongly maintained throughout evolution. More importantly, none of the amino acids described as important for this function were changed (apart from punctual changes) during EIV NS1 evolution, i.e. R38, K41, P31, D34, R35, R46, E97, E196, T5, G45, T49) (Wang et al., 1999, Talon et al., 2000, Pichlmair et al., 2006, Gack et al., 2009, Rajsbaum et al., 2012, Kuo et al., 2010). In addition, residues important for PKR inhibition were conserved during evolution, with some rare exceptions. This suggests that maintaining an NS1 protein with a strong control over IFN-β induction was selectively advantageous for EIV.

Surprisingly, no specific control of IFN-β induction could be detected for the C-terminal truncation, while it seemed to affect NS1 control over general gene expression and ISG induction (Figure 4-18-B & 4-18-C). This result in contradiction with previous reports from the literature showing that ponies experimentally infected with an EIV from 2003, possessing a truncated NS1 protein, were producing higher levels of type I IFN than ponies infected with earlier strains containing a full length NS1 protein (Daly et al., 2011, Wattrang et al., 2003). Similar results were found with swine and turkey influenza viruses (Cauthen et al., 2007, Solorzano et al., 2005). This discrepancy could be due to the difference in methodology, or perhaps some differences could not be detected in an artificial reporter assay done in human HEK293T cells.

Finally, residues controlling NS1 subcellular localization or important for the protein post-translational modification or activation of the PI3K pathway were in vast majority conserved. Thus, these functions may be important for EIV NS1. In addition, this implies that we would not expect change in NS1 subcellular localization upon infection.

Overall, the results of this chapter suggest that an emergent IAV relies on an NS1 protein with strong abilities to repress general gene expression after inter-species jump, but as the virus evolves in its new host, it accumulates mutations in the NS1 gene and the ones that target specifically the cellular anti-viral response are positively selected, while the ones controlling shutoff of cellular gene expression are preferentially eliminated.

Chapter 5

Impact of NS1 evolution on EIV infection phenotype

5. Impact of NS1 evolution on EIV infection phenotype

5.1. Introduction

Upon influenza virus infection, a variety of intracellular signalling pathways and vital cellular processes are activated. They are in part required to mount an antiviral response to infection, but they can also be exploited by the virus to support its replication (Ludwig et al., 2003, Ludwig et al., 2006). For example, the translation machinery remains competent for protein synthesis during an IAV infection and functional cellular mRNAs are still present in the cytoplasm of infected cells, however, the virus causes a switch from cellular to viral protein synthesis (Katze and Krug, 1984), and viral mRNAs are preferentially translated at the expense of cellular mRNAs (Garfinkel and Katze, 1993). The NS1 protein, which is expressed at high levels in infected cells (Krug and Etkind, 1973, Hutchinson et al., 2014), plays a major role in this function. It has notably been described to enhance viral translation (de la Luna et al., 1995, Enami et al., 1994), and alter cellular gene expression during infection (including those of type I IFN and ISGs) by interfering with the function of CPSF30 (Nemeroff et al., 1998) and PABPII (Chen et al., 1999). NS1 has also been shown to act as a type I IFN antagonist at different stages during viral infection (Garcia-Sastre et al., 1998) by blocking IRF-3, NF-κB, and c-Jun/ATF2 activation (Ludwig et al., 2002, Talon et al., 2000a, Wang et al., 2000), by repressing RIG-I (Mibayashi et al., 2007, Opitz et al., 2007, Pichlmair et al., 2006) and IPS-1 (Iwai et al., 2010), by inhibiting STAT1 and STAT2 phosphorylation (Wang et al., 2017), or by interfering with the function of ISGs, such as PKR and the OAS/RNase L pathway (Hale et al., 2008, Min and Krug, 2006).

A more recent addition to NS1 functions is the activation of PI3K (Ehrhardt et al., 2006), presumably by association with its regulatory subunits p85β. In eukaryotic cells, PI3K controls various cellular processes, such as cell growth, proliferation, and survival (Brazil et al., 2004, Cantley, 2002, Franke et al., 2003), and is involved in dsRNA-induced activation of IRF3 in infected cells (Sarkar et al., 2004). In doing so, IAV exploits the PI3K/Akt pathway to regulate virus-supportive functions, notably at later stages of viral replication to confers an anti-apoptotic signal via activation of Akt (Ehrhardt et al., 2006). Thus, despite initial work attributing to NS1 the capacity to promote apoptosis in infected cells (Schultz-Cherry et al., 2001), NS1 would prevent the premature induction of apoptosis via activation of PI3K, hence ensuring efficient replication for longer period (Ehrhardt et al., 2007, Hale et al., 2006). Interestingly, an intact C-terminus downstream of position 125, as well as amino acids 181 to 185 seem to be important for this function (Ehrhardt et al., 2007).

NS1 is not absolutely essential for virus replication in cell culture, as viruses lacking the NS1 gene (delNS1) or carrying large C-terminal deletions of the protein are able to replicate efficiently in interferon incompetent systems (Bergmann et al., 2000, Garcia-Sastre et al., 1998, Kochs et al., 2007b, Talon et al., 2000b, Nogales et al., 2014, Egorov et al., 1998). However, these viruses have are strongly attenuated in normal tissue culture cells (Egorov et al., 1998, Garcia-Sastre et al., 1998), which highlight the importance of NS1 in virus-host cell interplay. Previous studies have also shown that the NS segment, encoding for the NS1 protein, impact on IAV pathogenesis (Vergara-Alert et al., 2014), and specific mutations in NS1 RBD, ED or PBM have been shown to affect viral replication (Ehrhardt et al., 2006, Ehrhardt et al., 2007, Hale et al., 2006, Shin et al., 2007a, Shin et al., 2007b, Shin et al., 2007c), and to modulate virulence in a strain- and host-dependent manner (Zielecki et al., 2010, Seo et al., 2002, Jiao et al., 2008, Li et al., 2006, Jackson et al., 2008, Dankar et al., 2011, Soubies et al., 2010). More importantly, as little as one amino acid change in NS1 can significantly affect viral infection phenotype (Jackson et al., 2008, DeDiego et al., 2016, Steidle et al., 2010).

A great deal of work has aimed at correlating strain-specific differences of NS1 function with infection phenotype and virulence (reviewed in (Hale et al., 2008, Ayllon and Garcia-Sastre, 2015)). However, which specific NS1 mutations are preferentially selected during viral adaptation to a mammal, and how they affect virus infection phenotype remain unclear.

Aims: The aims of this chapter were to confirm the role of the NS1 gene in EIV infection phenotype; to evaluate the impact of NS1 evolutionary markers on EIV infectivity in mammalian cells; and to identify which advantages are associated with NS1 evolutionary markers E186K and C-terminal truncation.

5.2. Results

5.2.1. *In vitro* characterization of U/63 and O/03-WT and -NS reassortant viruses is mammalian cell lines

In order to determine whether the NS segment, which encodes for the NS1 protein, was a determinant of EIV infection phenotype, an eight plasmid-based reverse genetics system was used to generate U/63 and O/03 viruses, as well as NS segment reassortant of these two viruses (U/63-NS O/03 and O/03-NS U/63), as described in Chapter 3, section 3.5. The genotype of each virus was checked after viral rescue as described in Chapter 3.5.4 (Bavagnoli et al., 2015). A comparison of the replication efficiency of all viruses was carried out in MDCK and E.Derm cells, and their ability to spread between neighbouring cells (Hossain et al., 2008, Na et al., 2016) was also evaluated.

5.2.1.1. Growth kinetics of U/63 and O/03-WT and -NS reassortant viruses in MDCK cells

To compare the replication efficiency of U/63 and O/03 WT and NS reassortant viruses, MDCK cells were infected with each virus with 0.01 FFU/cell, or mock infected, and their growth kinetics was monitored over a 72h period.
As shown in (Figure 5-1), U/63 WT displayed a high replication efficiency early post-infection. Indeed, at 12 hpi the viral titre was already above 10^7 FFU/ml. U/63 maintained a similar titre for the following 36 h of infection (between 10^7 and 10^8 FFU/ml) after which it progressively declined to 10^6 FFU/ml at 72 hpi. In contrast, the growth rate of O/03 WT virus was slower, and increased progressively from 4 to 24 hpi, then peaked at 24 hpi at 10^9 FFU/ml and maintained a similar titre for the remaining of the monitored time period.

Interestingly, the introduction of O/03 NS segment into the U/63 backbone (U/63-NS-O/03) was associated with a reduction in viral growth rate and the maintenance of a high titre for a longer period of time (Figure 5-1). Indeed, at 12 hpi the U/63-NS-O/03 reassortant virus was two logs lower than its WT counterpart (10^5 FFU/ml versus 10^7 FFU/ml, respectively), but it reached a higher titre between 24 and 72 hpi (up to 2 logs higher than WT). In contrast, the introduction of the U/63 NS segment into the backbone of O/03 (O/03-NS-U/63) did not significantly affect O/03 viral growth, as O/03-NS-U/63 reassortant virus showed a similar growth pattern as its WT counterpart.

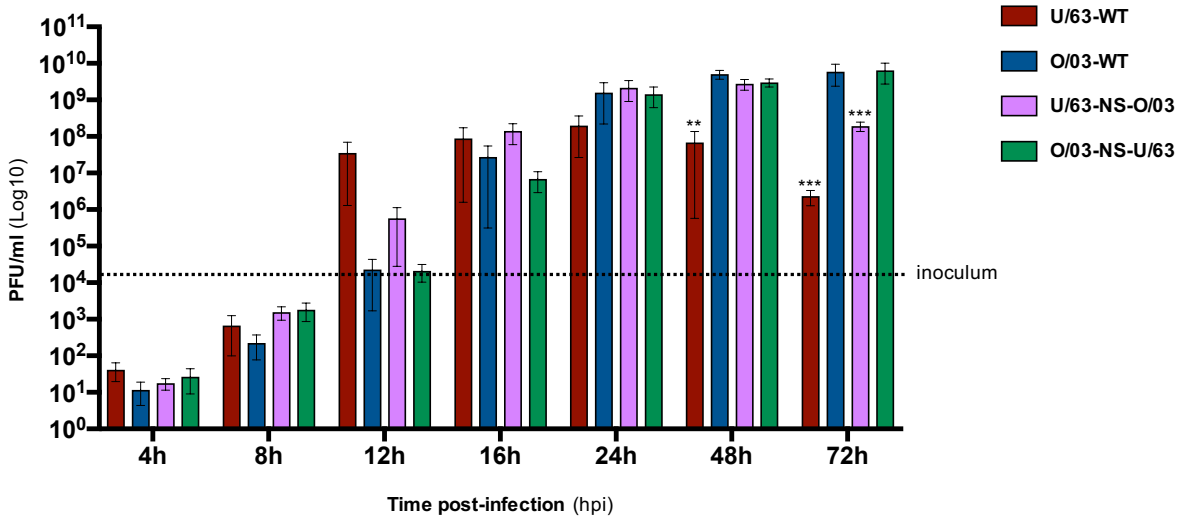


Figure 5-1: Growth kinetics of U/63 and O/03-WT and -NS swap mutants in MDCK cells.

MDCK cells were infected with U/63 and O/03-WT and -NS swap mutant viruses (0.01 FFU/cell), and viral growth kinetics were determined by measuring virus titres in supernatants at indicated times post-infection. Bars correspond to mean of three independent experiments and error bars represent SEM. Significance was calculated by Two-way ANOVA followed by Bonferroni's multiple comparisons test. ***, $p<0.001$ for U/63-WT and U/63-NS-O/03 versus O/03-WT and O/03-NS-U/63; **, $p<0.01$ for U/63-WT versus O/03-WT. The number of infectious particles inoculated (at 0h) is indicated by a dotted line.

5.2.1.2. Plaque phenotype of U/63 and O/03-WT and -NS reassortant viruses in MDCK cells

The ability of WT and NS reassortant mutant viruses to spread between neighbouring cells was then evaluated in MDCK cells at 48 hpi (Figure 5-2). It was observed that O/03-WT was able to spread from cell-to-cell more efficiently than the three other viruses. Indeed, the mean in O/03-WT plaque size was 12mm^2 in diameter compared to 3mm^2 for U/63-WT and both NS reassortant viruses.

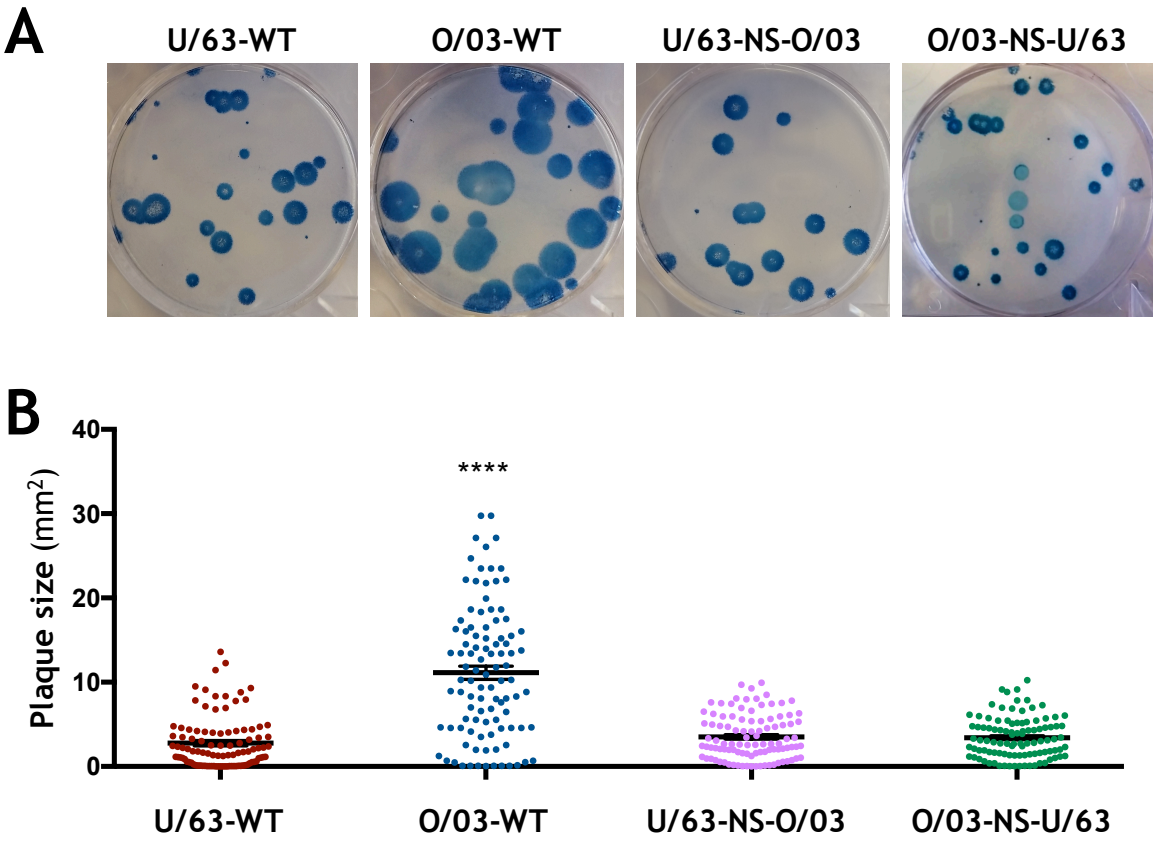


Figure 5-2: Plaque phenotype of U/63 and O/03-WT and -NS swap mutants in MDCK cells.

(A) Plaque phenotype of U/63 and O/03-WT and -NS swap mutant viruses were determined in MDCK cells at 48 hpi. A representative picture of three independent experiments is shown. **(B)** The average plaque size of each virus was measured by counting 100 plaques per conditions. Dots represent individual plaque size, bars correspond to mean, and error bars represent SEM. Significance was calculated as for Figure 5-1, with ****, $p<0.0001$ between O/03-WT and the three other viruses.

5.2.1.3. Growth kinetics of U/63 and O/03-WT and -NS reassortant viruses in E.Derm cells

When the replication efficiency of WT and NS reassortant viruses was compared in equine cells (Figure 5-3), it was noticed that O/03-WT grew to significantly higher titre than the three other viruses. This virus produced an increasing number of infectious particles from 4 to 24 hpi, and reached the infection peak at 24 hpi at 10^7 FFU/ml. This virus also maintained a titre above 10^6 FFU/ml for the 72 hpi monitored.

In contrast, U/63-WT only reached 10^4 FFU/ml (level of inoculum) at 24 hpi and maintained a similar titre for the remaining 48 hpi.

The introduction of O/03 NS segment in U/63 backbone did not affect the growth pattern of U/63, as previously observed in MDCK cells (Figure 5-1).

Interestingly, in these cells the introduction of U/63 NS segment in O/03 backbone strongly affected O/03 viral replication efficiency. This was particularly evident between 24 and 72 hpi, where the titre of O/03-NS-U/63 reassortant virus remained two logs lower than WT (between 10^5 FFU/ml and 10^4 FFU/ml).

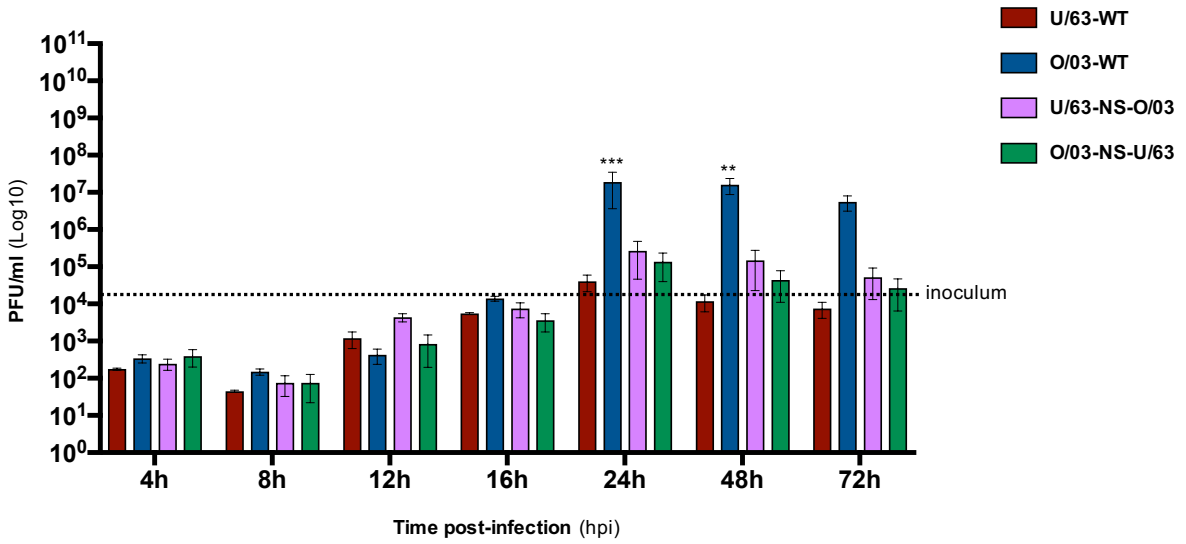


Figure 5-3: Growth kinetics of U/63 and O/03-WT and -NS swap mutants in E.Derm cells.

E.Derm cells were infected with U/63 and O/03-WT and -NS swap mutant viruses (0.1 FFU/cell), and viral growth kinetics were determined by measuring virus titres in the cell supernatant at indicated times post-infection. Bars correspond to mean of three independent experiments and error bars represent SEM. Significance was calculated as for Figure 5-1. ***, $p<0.001$ and **, $p<0.001$ for O/03-WT versus the three other viruses. The number of infectious particles inoculated (at 0h) is indicated by a dotted line.

Taken together these results confirmed that the NS segment, encoding for the NS1 protein, is a determinant of EIV infection phenotype. Furthermore, the effect of NS reassortment on viral growth kinetics and cell-to-cell spread seemed to be viral context- and cell type-dependent.

5.2.2. *In vitro* characterization of U/63 and O/03-WT and NS1 mutants in mammalian cells

5.2.2.1. Introduction of codon 186 and 220 mutations in U/63 and O/03 NS segment

The NS segment of U/63 and O/03 possess a total of thirty different nucleotides, which result in three different amino acids for NEP, and twelve for NS1. In addition, U/63

expresses a full-length NS1 (230 amino acids), while O/03 expresses a truncated version of the protein (219 amino acids) (Appendix 31). Thus, to discriminate between these differences, and evaluate the role of NS1 residue 186 and C-terminus on EIV infection phenotype, a more targeted approach was necessary. The use of a reverse genetic system in association with a site-directed mutagenesis approach allowed the generation of isogenic mutants of U/63 and O/03 for NS1 codon 186 and 220 (Chapter 3, section 3.1.8 and 3.5.3). Briefly, U/63-E186K-containing mutants (U/63-E186K and U/63-E186K-219) were generated by changing NS1 codon 186 for 'GAA' to 'AAA' in the NS segment (Figure 5-4, Appendix 32). In a similar fashion, U/63-219-containing mutants (U/63-219 and U/63-E186K-219) were generated by changing NS1 codon 220 from 'CGA' to 'TGA'. The opposite mutations were introduced in O/03 NS segment to generate O/03-K186E-containing mutants (O/03-K186E and O/03-K186E-230), and O/03-230 mutants (O/03-230 and O/03-K186E-230) (Figure 5-5, Appendix 33). None of the introduced mutations affected the NEP open reading frame (Appendices 34).

The growth kinetics and plaque phenotype of U/63 and O/03 WT and NS1 mutant viruses was then carried out in MDCK and equine cells.

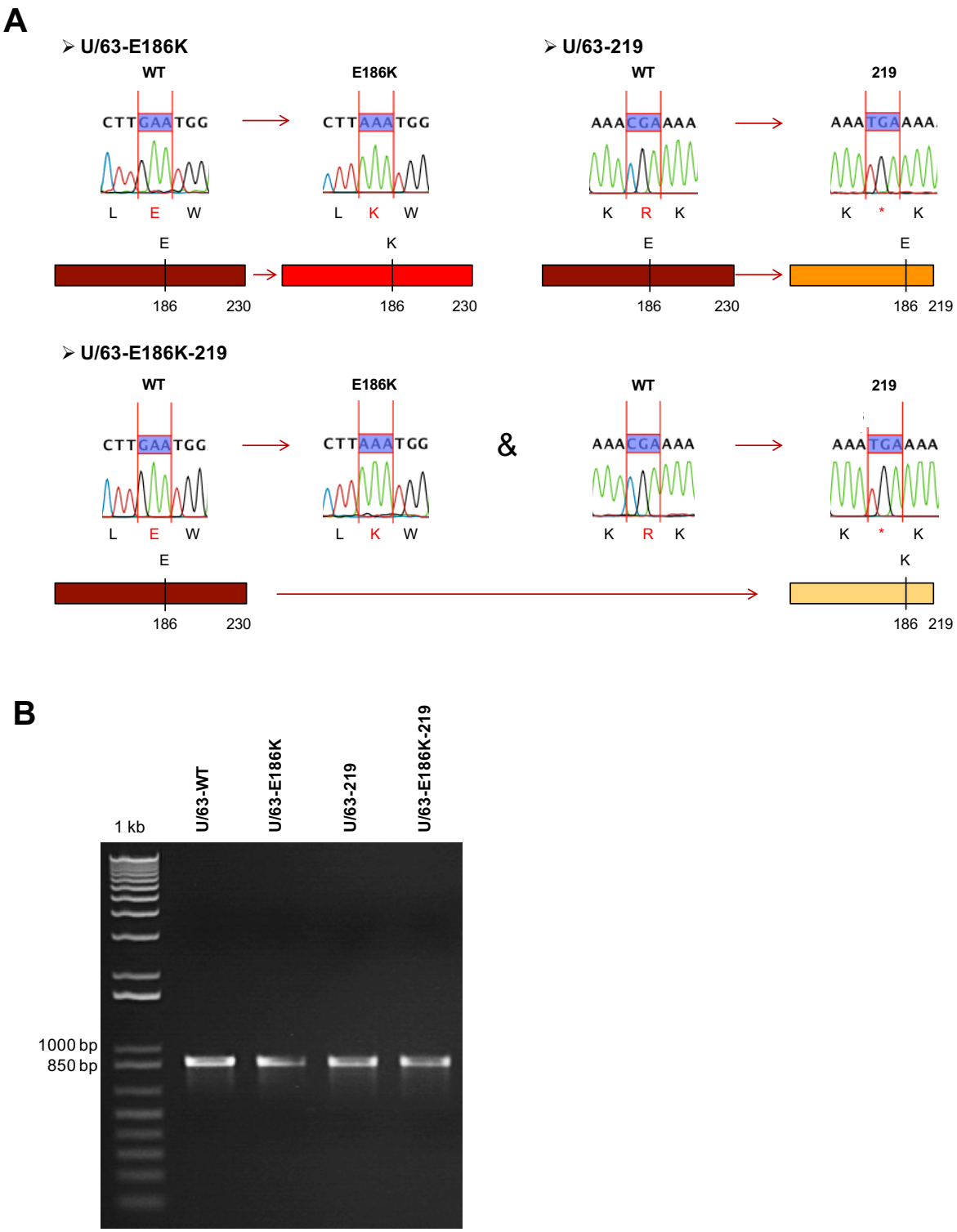


Figure 5-4: Generation of U/63 NS mutants for residue 186 and C-terminus.

In a similar fashion to Figure 4-10, **(A)** the NS segment of WT, single or double mutant viruses were sequenced by Sanger sequencing and the corresponding traces are shown. **(B)** The size of each segment (890bp) were checked.

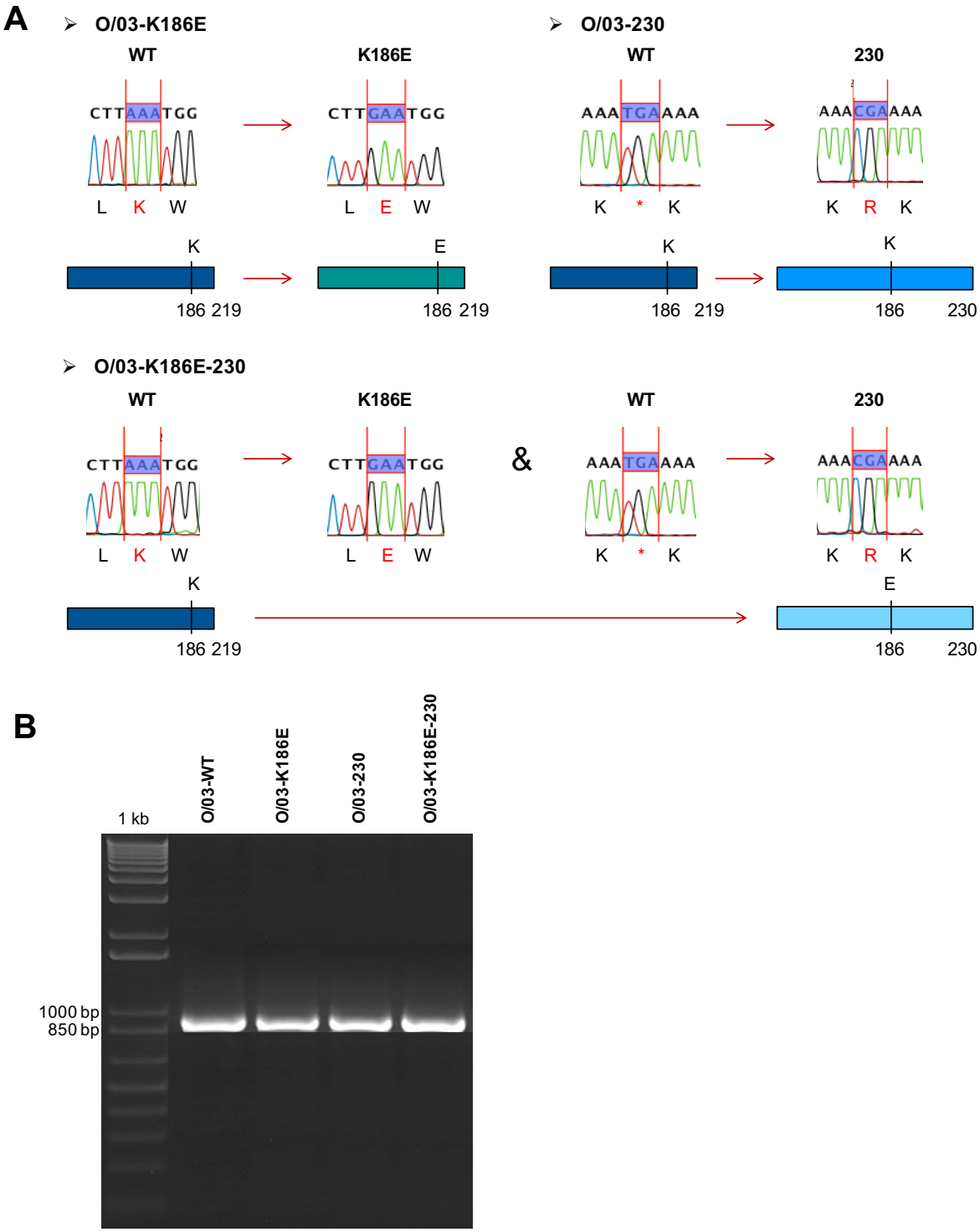


Figure 5-5: Generation of O/03 NS mutants for residue 186 and C-terminus.

In a similar fashion to Figure 4-10, **(A)** the NS segment of WT, single or double mutant viruses were sequenced by Sanger sequencing and the corresponding traces are shown. **(B)** The size of each segment (890bp) were checked.

5.2.2.2. Growth kinetics of U/63 and O/03-WT and mutants in MDCK cells

As shown in (Figure 5-6), U/63 NS1 single and double mutants showed differences in growth kinetics in MDCK cells compared to their WT counterpart. These differences were particularly marked at 12, 16, 48 and 72 hpi. Indeed, the titre of mutant viruses were 3 to 4 logs lower than WT at 12 and 16 hpi (10^3 and 10^5 FFU/ml for E186K mutants versus 10^7 and 10^8 FFU/ml for WT), while being 2 logs higher than WT between 48 and 72 hpi (10^9 and 10^8 FFU/ml for E186K mutants versus 10^7 and 10^6 FFU/ml for WT, respectively).

The titre of U/63 E186K-containing mutant viruses also peaked 24 h later than WT (48 versus 24 hpi, respectively), but the mutant viruses reached a higher titre at the peak of infection (10^9 FFU/ml for E186K-containing mutants versus 10^8 FFU/ml for WT).

The U/63-219 mutant showed an intermediate phenotype. Indeed, between 12 and 16 hpi U/63-219 titre was 2 logs higher than E186K-containing mutants (10^5 and 10^7 FFU/ml versus 10^3 and 10^5 FFU/ml, respectively), but remained 2 logs lower than WT (10^5 and 10^7 FFU/ml versus 10^7 and 10^8 FFU/ml, respectively). In addition, U/63-219 titre peaked 24 h later and 1 log higher than WT (10^9 FFU/ml versus 10^8 FFU/ml, respectively).

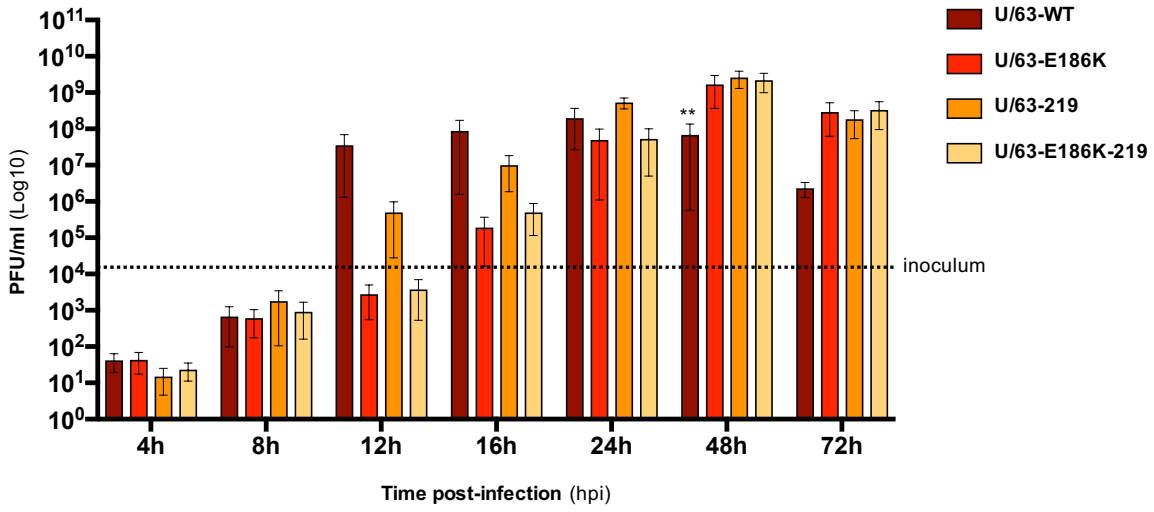


Figure 5-6: Growth kinetics of U/63-WT and mutant viruses in MDCK cells.

MDCK cells were infected with U/63-WT and mutant viruses (0.01 FFU/cell), and viral growth kinetics were determined as for Figure 5-2. Bars correspond to mean of three independent experiments and error bars represent SEM. Significance was calculated as in Figure 5-1, and **, $p<0.01$ for U/63-WT versus the three other viruses. The number of infectious particles inoculated (at 0h) is indicated by a dotted line.

When the growth rate of O/03-WT and mutant viruses was compared, different results were obtained. Indeed, the single introduction of E186K substitution or C-terminal extension in NS1 did not significantly affect O/03 replication efficiency in MDCK cells (Figure 5-7). However, when introduced together, NS1 K186E and C-terminal extension strongly modified the replication pattern of O/03. The double mutant virus (O/03-K186E-230) grew faster than WT and reached a higher titre than any other viruses at 12 hpi (10^7 FFU/ml). Then, the virus titre plateaued between 12 and 72 hpi (at 10^7 and 10^8 FFU/ml). Interestingly, this virus never reached a titre as high as the three other viruses during the entire course of infection.

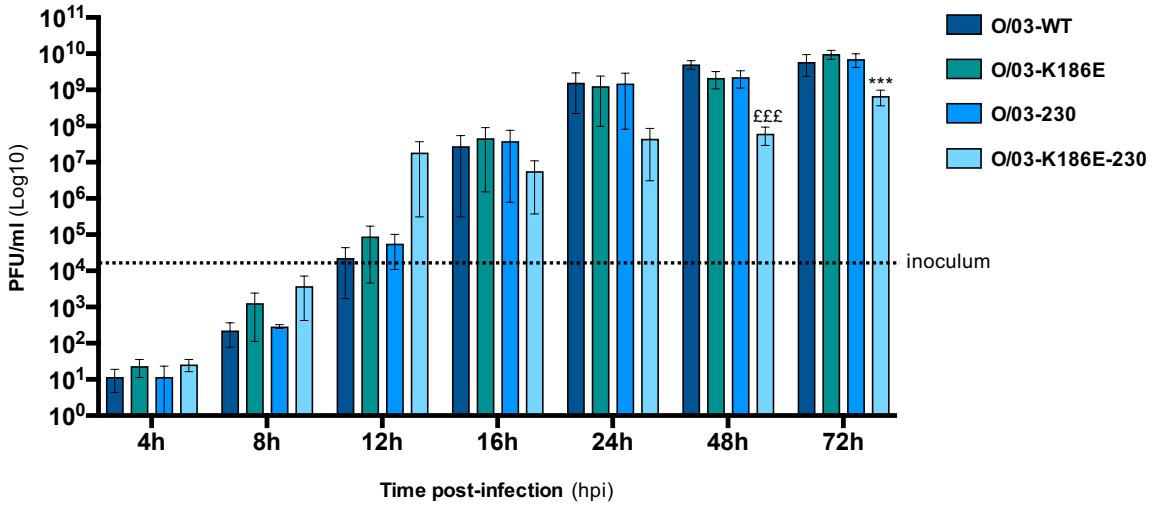


Figure 5-7: Growth kinetics of O/03-WT and mutant viruses in MDCK cells.

MDCK cells were infected with O/03-WT and mutant viruses (0.01 FFU/cell), and viral growth kinetics were determined as for Figure 5-2. Bars correspond to mean of three independent experiments and error bars represent SEM. Significance was obtained as for Figure 5-1. ***, $p<0.001$ for O/03-K186E-230 versus all the three other viruses, and £££, $p<0.001$ for O/03-K186E-230 versus O/03-WT. The number of infectious particles inoculated (at 0h) is indicated by a dotted line.

5.2.2.3. Plaque phenotype of U/63 and O/03-WT and mutants in MDCK cells

The capacity of isogenic mutants to spread between neighbouring cells was then compared to WT. As shown below, U/63 mutant viruses displayed similar plaque phenotypes than WT (Figure 5-8).

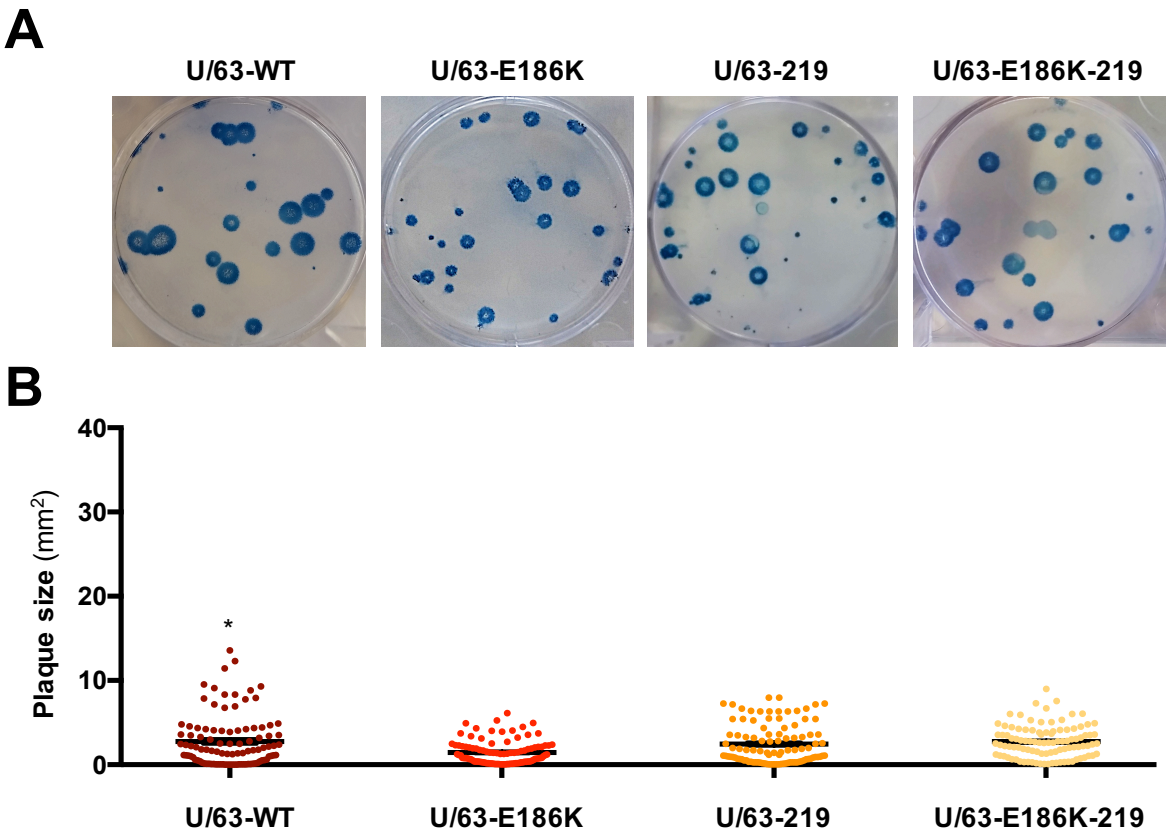


Figure 5-8: Plaque phenotype of U/63-WT and mutant viruses in MDCK cells.

(A) Plaque phenotype of U/63-WT and mutant viruses was determined in MDCK cells at 48 hpi. A representative picture of three independent experiments is shown. **(B)** The average plaque size of each virus was measured as for Figure 5-3. Dots represent individual plaque size, bars correspond to mean, and error bars represent SEM. Significance was calculated as for Figure 5-1, *, $p<0.05$ for U/63-WT versus the three mutant viruses.

However, in the case of O/03, K186E substitution and/or C-terminal extension affected cell-to-cell spread in MDCK cells (Figure 5-9). Indeed, all mutant viruses displayed smaller plaque sizes than WT, and the most dramatic effect was observed with the double mutant virus (O/03-K186E-230).

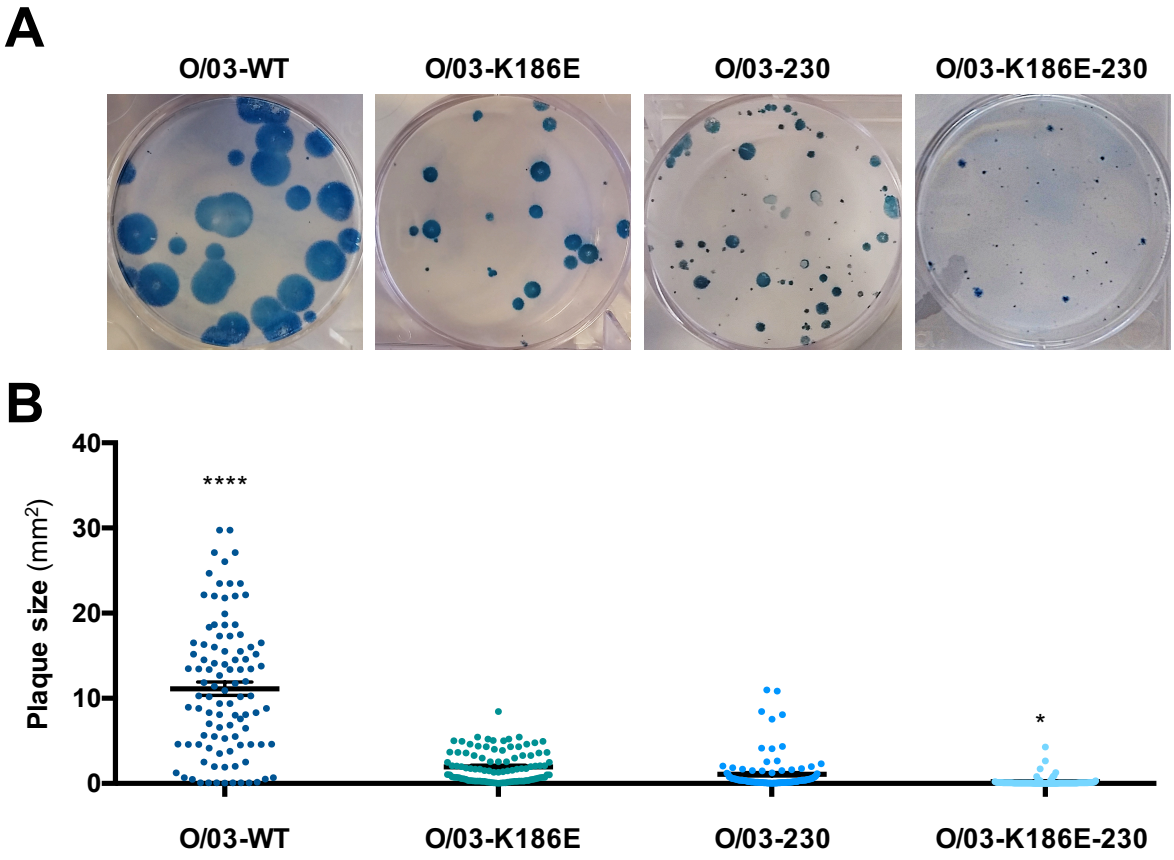


Figure 5-9: Plaque phenotype of O/03-WT and mutant viruses in MDCK cells.

(A) Plaque phenotype of O/03-WT and mutant viruses was determined in MDCK cells at 48 hpi. A representative picture of three independent experiments is shown. **(B)** The average plaque size of each virus was measured as for Figure 5-3. Dots represent individual plaque size, bars correspond to mean, and error bars represent SEM. Significance was calculated as for Figure 5-1. ****, $p<0.0001$ for O/03-WT versus all other viruses; *, $p<0.05$ for O/03-K186E-230 versus O/03-K186E.

5.2.2.4. Growth kinetics of U/63 and O/03-WT and mutants in E.Derm cells

When growth of WT and mutant viruses was compared in E.Derm cells (Figure 5-10), it was observed that E186K substitution and/or C-terminal truncation of NS1 of U/63 did not improve viral growth rate, and after 24 hpi, it even seemed to reduce it as all mutant virus titres were 1 to 2 logs lower than WT (between 10^3 and 10^2 FFU/ml for mutant viruses versus 10^4 FFU/ml for WT).

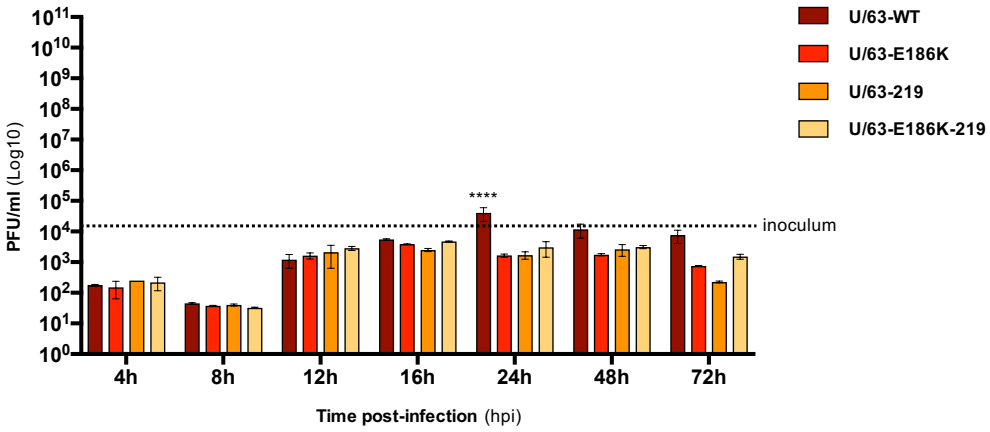


Figure 5-10: Growth kinetics of U/63-WT and mutant viruses in E.Derm cells.

E.Derm cells were infected with U/63-WT and mutant viruses (0.1 FFU/cell), and viral growth kinetics were determined as for Figure 5-4. Bars correspond to mean of three independent experiments and error bars represent SEM. Significance was calculated as for Figure 5-1. ****, p<0.0001 for U/63-WT versus the mutant viruses. The number of infectious particles inoculated (at 0h) is indicated by a dotted line.

In the context of O/03, all mutations introduced affected viral replication efficiency (Figure 5-11). The most dramatic effect was observed for O/03-E186K single mutant, which did not seem to grow (titres below 10^2 FFU/ml for the 72h of infection). In addition, viruses expressing full length NS1 proteins (O/03-230 and O/03-K186E-230) grew to similar titres than WT during the first 16 hpi (from 10^2 to 10^4 FFU/ml), after which their titres declined (10^4 to 10^2 FFU/ml).

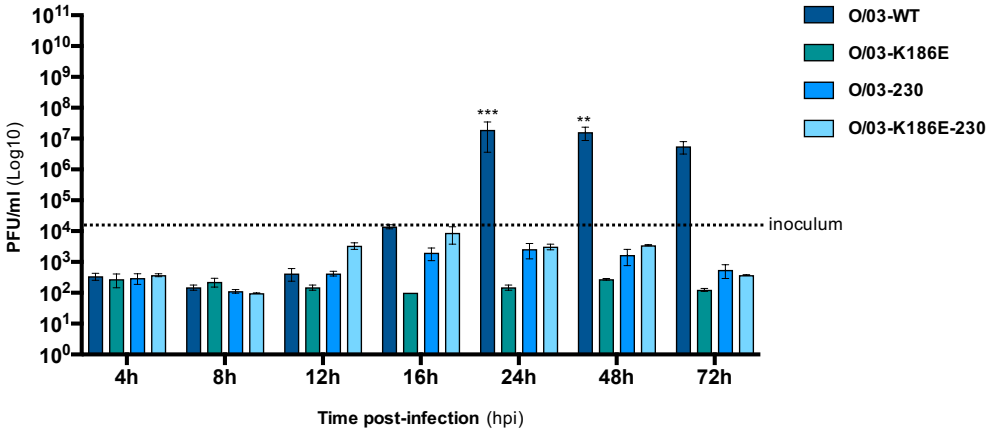


Figure 5-11: Growth kinetics of O/03-WT and mutant viruses in E.Derm cells.

E.Derm cells were infected with O/03-WT and mutant viruses (0.1 FFU/cell), and viral growth kinetics were determined as for Figure 5-4. Bars correspond to mean of three independent experiments and error bars represent SEM. Significance was obtained as for Figure 5-1, ***, p<0.001 and **, p<0.01 between O/03-WT and the three other viruses. The number of infectious particles inoculated (at 0h) is indicated by a dotted line.

Taken together these results confirmed the importance of NS1 and the nature of residue 186 and the C-terminal tail length on EIV infection phenotype. Moreover, they confirmed that NS1 impact on EIV growth kinetics and plaque phenotype in a different manner depending on the viral background and the cell type used.

5.2.3. Impact of K186E and C-terminal extension on EIV control of infected cells

To study in more details the role of NS1 amino acid 186 and C-terminal tail on virus-host interaction, O/03 viruses were selected. First the ability of O/03-WT and mutant viruses to shutdown total protein production upon infection was evaluated. Protein shutdown upon infection was also compared to viral protein production, total antiviral cytokines production in cell supernatant, as well as expression of ISGs and induction of apoptosis during infection in equine cells. Finally, the replication efficiency of these four viruses was compared in equine cells either pre-exposed to type I IFN or after having blocked the cellular response to type I IFN by a chemical approach.

5.2.3.1. Shutdown of general protein production upon infection

To determine the effects of NS1 K186E and C-terminal extension on O/03 ability to control cellular protein synthesis, puromycin assays were performed at various times post-infection in equine cells, as described in Chapter 3, section 3.6. Under appropriate conditions, puromycin can be incorporated at the C-terminus of all nascent proteins in eukaryotic cells ((Tabuchi, 2003)). After cell lysis, an immunoblotting against puromycin can performed and allows the detection of nascent proteins that were being produced at the moment of cell lysis. Total protein production will appear as a black smear, and the darker and longer the smear is, the more proteins were being produced at the moment of cell lysis.

As shown in (Figure 5-12), only the double mutant virus induced a strong protein shutoff upon infection. This was particularly marked at 8 and 12 hpi. While this result is consistent with E186 playing an important role in blocking general gene expression via a CPSF30-dependent mechanism (Chapter 4), it also suggests that a complementary role of NS1 C-terminus might be necessary to induce a strong protein shutoff in equine cells, as O/03-K186E did not strongly block protein synthesis at any time post infection. However, we cannot rule out the possibility that O/03-K186E's lack of protein shutoff was due to a low

level of infection or expression of NS1, as this virus was strongly attenuated in equine cells. Interestingly, O/03-WT induced a significant increase in protein synthesis at 12 hpi, and a low but significant protein shutoff at 24 hpi

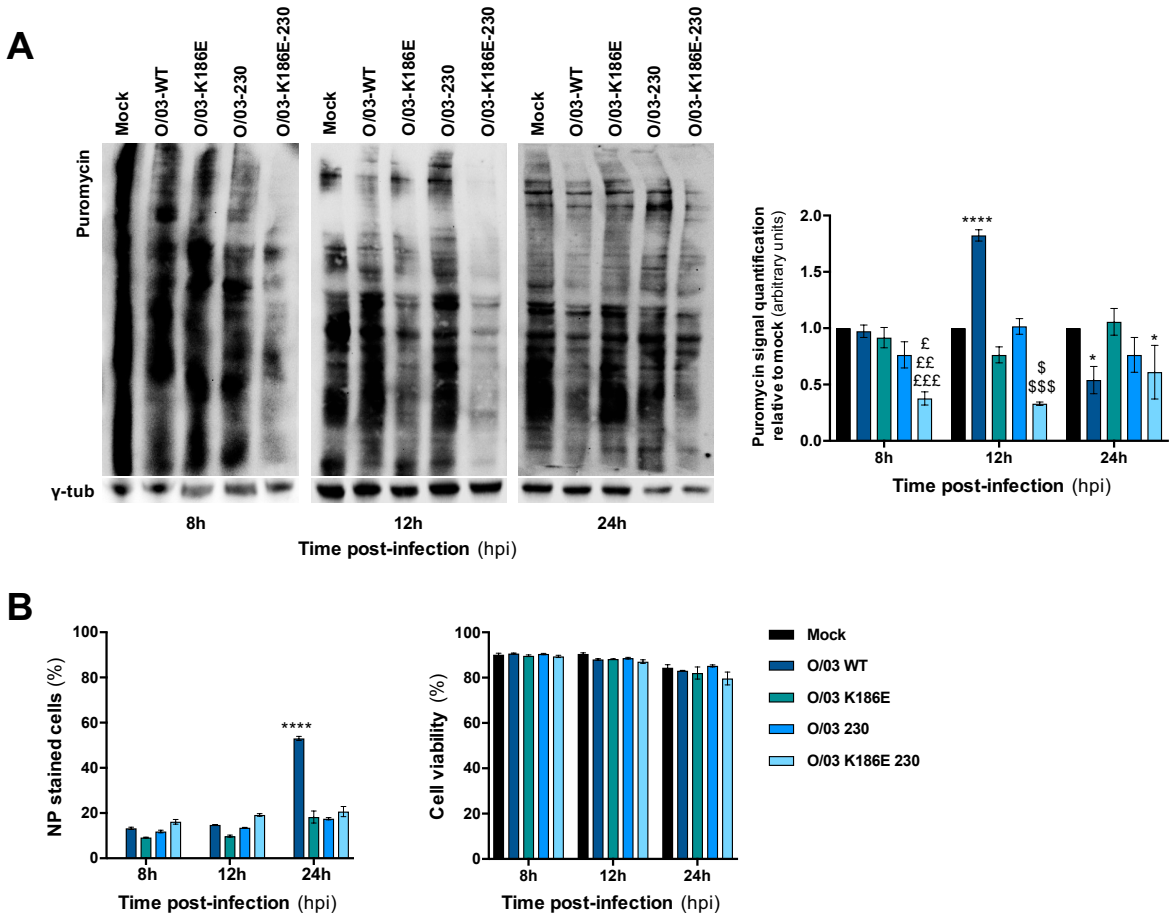


Figure 5-12: Protein shutoff in E.Derm cells upon infection with O/03 WT and mutant viruses.

E.Derm cells were infected with O/03-WT and mutant viruses (0.1 FFU/cell). **(A)** At the indicated time points, culture media were replaced with a puromycin-containing suspension and incubated for 1h. The cells were then lysed and immunoblotted for puromycin, as described in the methods section. Gamma tubulin was used as a loading control. The experiment was repeated three times independently and a representative picture of each western blot is shown. The quantification of the puromycin signal expressed relatively to mock has been obtained for three independent experiments, as described in the method section. Significance was obtained as in Figure 5-1. At 8 hpi: £££, $p<0.001$ for O/03-K186E-230 versus mock and O/03-WT; ££, $p<0.01$ for O/03-K186E-230 versus O/03-K186E; £, $p<0.05$ for O/03-K186E-230 versus O/03-230. At 12 hpi: ***, $p<0.001$ for O/03-WT versus all other conditions. $, $p<0.05$ for O/03-K186E-230 versus O/03-K186E and $$$, $p<0.001$ for O/03-K186E-230 versus O/03-230 and mock, respectively. At 24 hpi: *, $p<0.05$ for O/03-WT and O/03-K186E-230 versus mock and O/03-K186E. **(B)** Cells of parallel samples were collected, fixed in formalin and immunostained for the NP viral protein, as described in the method section. The percentage of infected cells and cell viability for each condition was then determined by flow cytometry. Significance was obtained as previously described and ****, $p<0.0001$ for O/03-WT versus all other conditions. Bars correspond to mean of three independent experiments and error bars represent SEM.

To rule out the possibility that differences in protein shutoff were due to different levels of infection, the number of E.Derm cells expressing the viral nucleoprotein (NP) was analysed by flow cytometry at 8, 12 and 24 hpi, as described in Chapter 3, section 3.5.7. As seen in (Figure 5-12-B), a comparable percentage of cells were infected with WT and mutant viruses at 8 and 12 hpi. At 24 hpi, a greater proportion of cells were expressing the viral NP upon O/03-WT infection, which confirms that this virus was more infectious than mutants in E.Derm cells (Figure 5-11).

5.2.3.2. Cellular response to viral infection

To compare WT and mutant virus abilities to limit the establishment of a cellular antiviral state upon infection, E.Derm cells were infected (0.1 FFU/cell) with WT and mutant viruses, or mock infected, and the level of expression of ISG15 and Mx1 was evaluated, as described in Chapter 3, section 3.8. Since the establishment of an antiviral state is associated with induction of apoptosis, the expression of both total and cleaved caspase 3 was evaluated by western blot at different times post infection. In addition, the level of expression of the NP and NS1 proteins were also monitored.

Interestingly, ISG15 and Mx1 were detected in cell lysates upon infection with both O/03-230 and O/03-K186E-230 mutant viruses, but not WT (Figure 5-13) at 24 and 48 hpi. Surprisingly, these proteins were not detected in cells infected with O/03-K186E.

When the level of apoptosis was compared in WT and mutant virus-infected cells, it was observed that cleaved caspase 3 was already detectable at 24 hpi in cells infected with the double mutant of O/03 as well as with O/03-230. Instead, apoptosis induction was detectable 24 h later during O/03-WT infection. Of note, no apoptosis induction upon O/03-K186E mutant infection was detected.

High levels of NS1 and NP proteins were also detected in cells infected with O/03-double mutant between 8 and 24 hpi. In O/03-WT-infected cells NP was detected at 24 and 48 hpi, and NS1 could only be detected at 48 hpi. Of note, no viral protein could be detected for O/03-K186E single mutant virus.

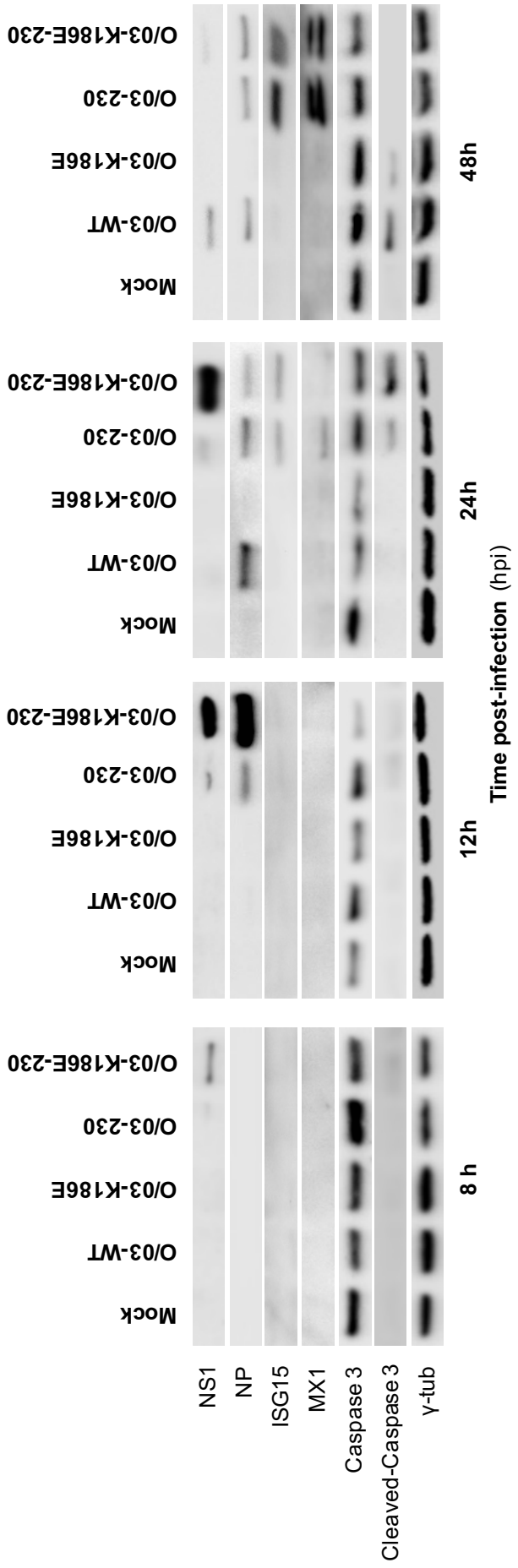


Figure 5-13: Control of apoptosis and ISG induction in E.Derm cells upon infection with O/03-WT and mutant viruses.

E.Derm cells were infected with O/03-WT and mutant viruses (MOI 0.1). At the indicated time points, cells were lysed and immunoblotted for the indicated proteins. The experiment was repeated three times independently and a representative picture of each western blot is shown.

5.2.3.3. Production of antiviral cytokines upon infection

To assess whether the IFN system was the limiting factor of O/03 mutant virus replication, the antiviral activity of the supernatant of equine cells infected with the different viruses were measured at several time points post infection using a VSV-GFP-based bioassay (Martinez-Sobrido et al., 2006) (Chapter 3, section 3.6). Antiviral cytokine production was also compared with viral growth kinetics (Figure 5-14).

Surprisingly, the highest level of antiviral cytokines was detected upon O/03-WT infection, peaking at 24 hpi (Figure 5-14-A). This did not seem to affect virus growth since this virus reached rapidly a very high titre (10^7 FFU/ml at 24 hpi).

The pattern of antiviral cytokine production upon O/03-230 and O/03-K186E-230 infection (Figure 5-14-C & -D) were very similar. Indeed, in both cases low levels of antiviral cytokines were produced, and this was associated with viral growth, although at a much lower level than O/03-WT. In contrast, in cells infected with O/03-K186E mutant virus, a relatively high peak of antiviral cytokines was detected at 8 hpi, and no viral growth was measured (Figure 5-14-B).

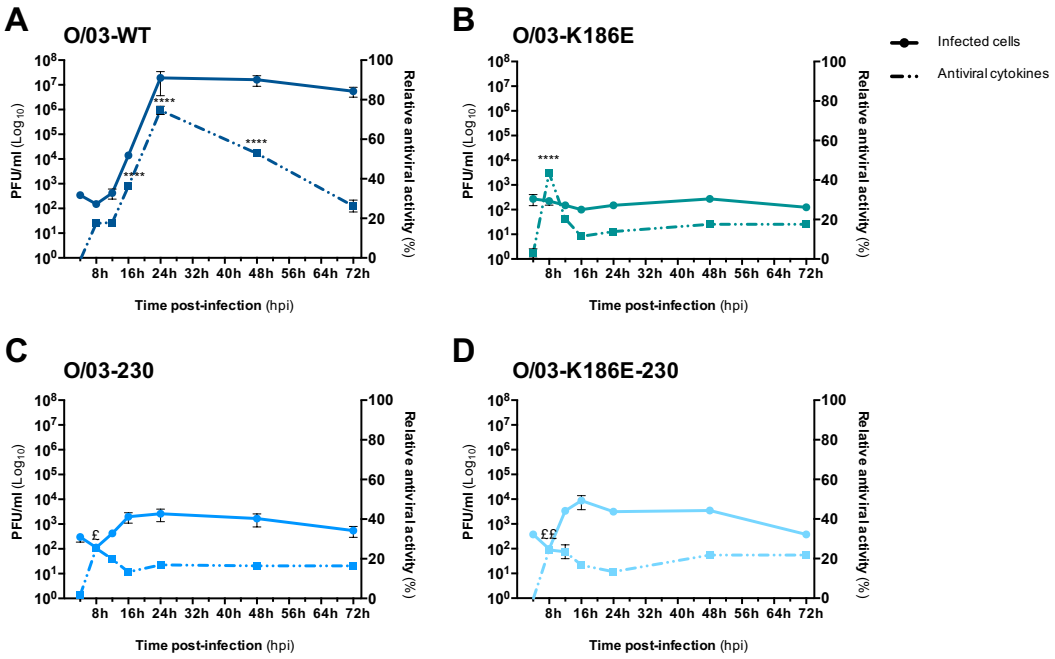


Figure 5-14: Comparison of growth kinetics and production of antiviral cytokine in equine cells infected with O/03 WT and mutant viruses.

E.Derm cells were infected (0.1 FFU/cell) for 72h with **(A) O/03-**WT, **(B) -**K186E, **(C) -**230, and **(D) -** K186E-230. Supernatants were collected, and cells fixed at indicated times pi. The total amount of antiviral cytokines produced upon infection was measured using an IFN bioassay and virus growth was determined by immunofocus assay. Error bars represent SEM. Significance was calculated as for Figure 5-1. ****, $p<0.0001$ for antiviral cytokine production upon infection with indicated viruses at indicated time pi compared to all other conditions. £, $p<0.05$ and ££, $p<0.01$ for antiviral cytokine production upon infection with indicated viruses at indicated time pi compared to WT.

5.2.4. Evaluation of WT and mutant virus replication efficiency after blockade of the cellular response to type I IFN

To confirm the role of the IFN pathway in limiting O/03 NS1 mutant replication, E.Derm cells were treated with Ruxolitinib (JAK inhibitor) for 24h prior to infection (Stewart et al., 2014) (Chapter 3, section 3.5.8). The number of infected cells (Figure 5-15-A) and the growth kinetics of each virus in the presence of Ruxolitinib (Figure 5-15-C) were monitored.

Interestingly, in the presence of Ruxolitinib, all viruses replicated to similar titres, and infected a growing number of cells. Surprisingly, the number of cells infected with the double mutant virus was significantly higher at early time points post infection compared to the three other viruses. These results confirmed the central role of the type I IFN pathway in limiting viral replication of mutant viruses in equine cells.

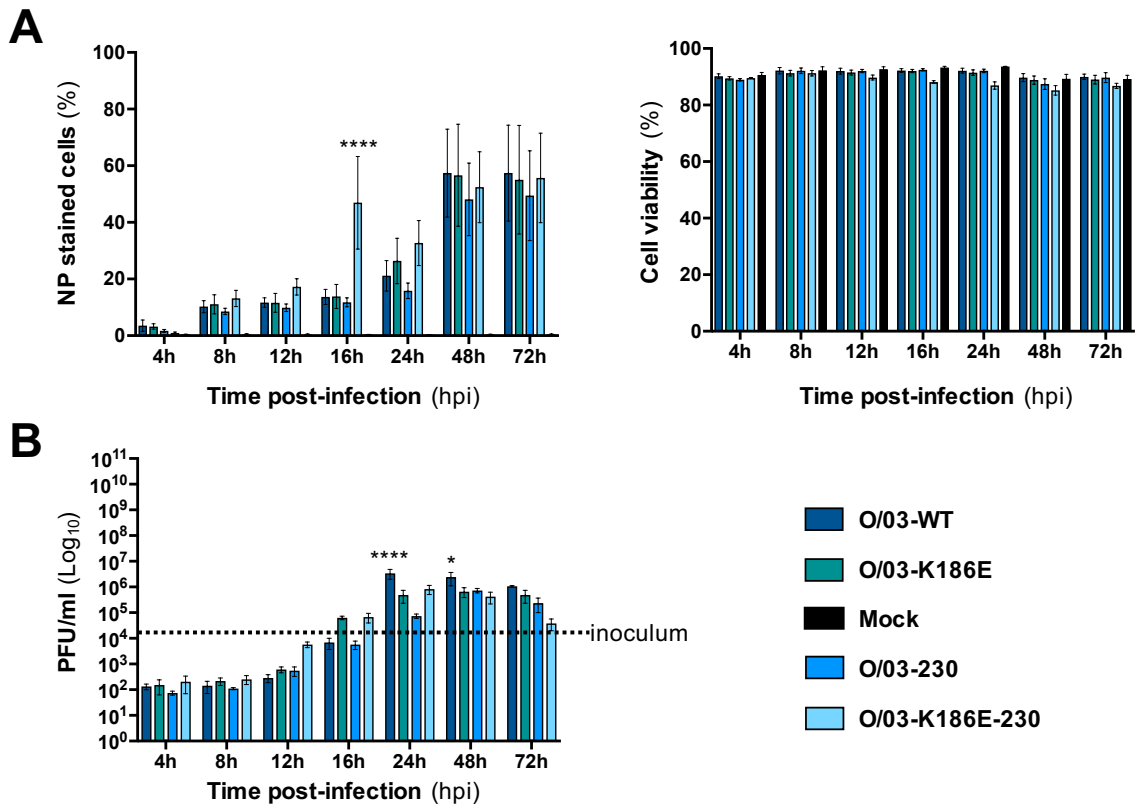


Figure 5-15: JAK inhibition restores growth kinetics of O/03 NS1 mutant viruses.

E-derm cells were treated for 24h with 4uM Ruxolitinib prior to infection (0.1 FFU/cell) with O/03-WT, -K186E, -230, and -K186E-230. Cells and supernatants were collected at different times post-infection. **(A)** Percentage of infected cells and cell viability was determined by flow cytometry. **(B)** Viral titers were measured as described in Methods. Bars correspond to mean of three independent experiments and error bars represent SEM. Significance was calculated as for Figure 5-1. ****, $p<0.0001$ for indicated virus at indicated time post infection versus all other conditions at the same time post infection.

5.2.5. Evaluation WT and mutant virus replication efficiency in type I IFN primed-equine cells

To confirm the importance of NS1 evolutionary markers K186 and C-terminal truncation, for efficient viral replication in the presence of type I IFN, E.Derm cells were pre-treated with Universal IFN for 24 h prior to infection (0.1 FFU/cell). As done previously, viral replication and percentage of infected cells were monitored for 72 h (Chapter 3, section 3.5.7 and 3.5.8).

While all viruses exhibited a reduction in viral growth (Figure 5-16-A), only the WT virus managed to infect an increasing number of cells (Figure 5-16-A) and produce an increasing quantity of infectious particles (Figure 5-16-B) in IFN-primed cells.

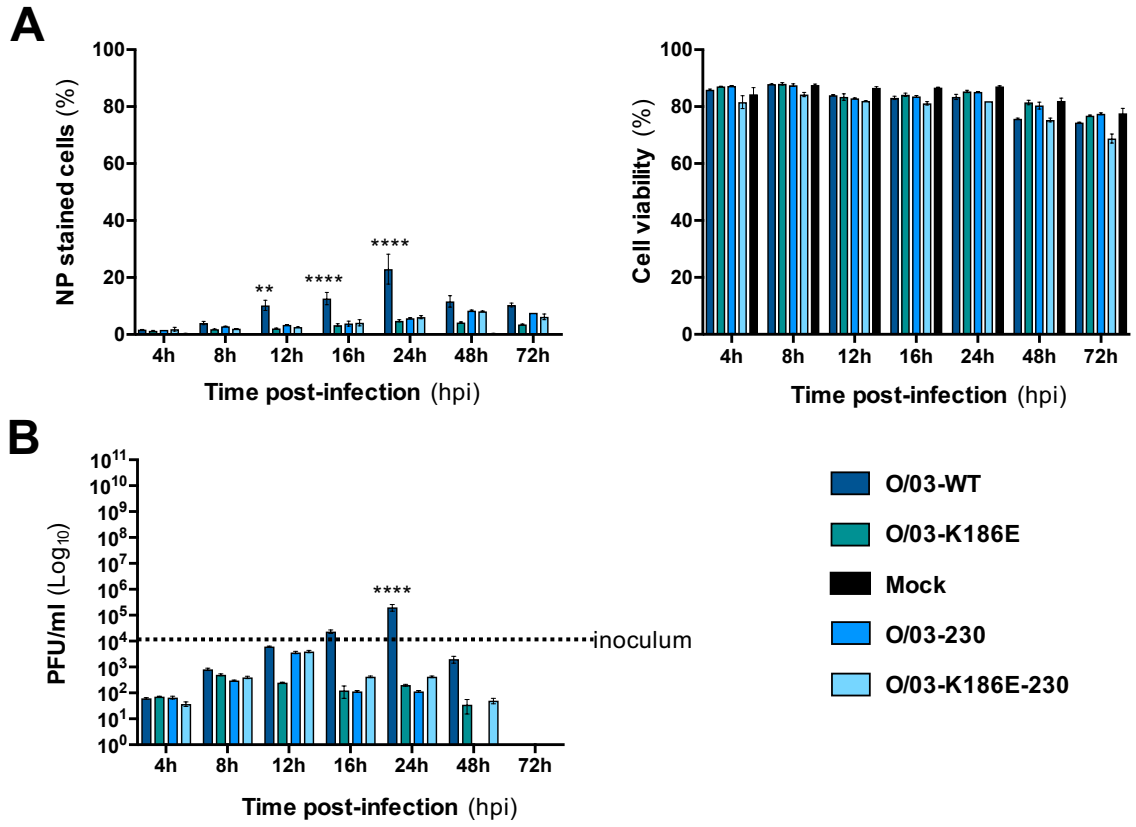


Figure 5-16: Growth kinetics of O/03 WT and NS1 mutant viruses in universal IFN-primed equine cells.

E.Derm cells were treated for 24h with universal IFN prior to infection (0.1 FFU/cell) with O/03-WT, -K186E, -230, and -K186E-230. Cells and supernatants were collected at different times post-infection. **(A)** Percentage of infected cells and cell viability was determined by flow cytometry. **(B)** Viral titres were measured as described in Methods. Bars correspond to mean of three independent experiments and error bars represent SEM. Significance was calculated as for Figure 5-1. ****, $p<0.0001$ for indicated virus at indicated time post infection versus all other conditions at the same time post infection.

5.3. Discussion

In this chapter, the effect of NS1 on EIV infection phenotype was evaluated. First, it was demonstrated that the NS segment, encoding the NS1 protein, affected EIV replication efficiency and cell-to-cell spread in mammalian cells. To do so, viral growth kinetics and plaque phenotype of WT and NS reassortant of U/63 and O/03 viruses were compared in MDCK and E.Derm cells. It was shown that the WT viruses displayed differences in growth kinetics (Figure 5-1 & -3) in both cell lines and plaque phenotype in MDCK cell (Figure 5-2). Although U/63-WT grew poorly in E.Derm cells (Figure 5-3), it replicated at high rate in MDCK cells and reached its peak of infection earlier (although at a lower titre) than O/03-WT (Figure 5-1). In contrast, O/03-WT grew more progressively and reached high titres in both cell lines (Figure 5-1 & -3). In addition, O/03-WT displayed larger plaques than U/63-WT in MDCK cells (Figure 5-2). Interestingly, reassortment of U/63 and O/03 NS segments modified these infection phenotypes. Furthermore, the effect was dependent on the viral context- and the cell type. Indeed, introduction of O/03 NS segment into U/63 backbone only changed U/63 growth pattern in MDCK cells (Figure 5-1), while its plaque phenotype (Figure 5-2) and growth kinetics were mostly unchanged in E.Derm cells (Figure 5-3). In contrast, O/03 growth pattern in E.Derm cells (Figure 5-3) and plaque phenotype in MDCK cells (Figure 5-2) were strongly affected by the NS segment reassortment, while its growth kinetic in MDCK cells was maintained (Figure 5-1). These results are in accordance with Shelton et al.'s work, who studied the replication capacity of NS reassortants between the human 2009pH1N1 virus and a human seasonal H3N2 virus, two viruses that express a short non-CPSF30 binder NS1 protein and a full-length NS1 that interacts with CPSF30, respectively (Shelton et al., 2012). They showed that the introduction of the H3N2 NS segment into the p2009pH1N1 virus reduced the replication efficiency of the latter in an IFN competent model. In addition, Twu and colleagues showed that the replication efficiency and plaque phenotype of an IAV expressing an NS1 protein that binds CPSF30 was strongly affected in MDCK cells when this binding was altered (Twu et al., 2006). Other reports have also described that for other IAVs, the two NS1 proteins could be swapped without affecting viral growth between (Kim et al., 2014). This further reinforce the hypothesis of species-specific functions of NS1.

To study in more details the effect of NS1 evolutionary markers on EIV infection phenotype, three isogenic mutants for U/63 and O/03 viruses carrying a mutation at codon 186 and/or codon 220 (Figure 5-4 & -5) in their respective NS segment –were generated. By comparing the replication efficiency and plaque phenotype (Figure 5-8 & -9) of mutant and WT viruses in MDCK (Figure 5-6 & -7) and E.Derm cells (Figure 5-10 & -11), it was confirmed

that the introduced mutation affected viral growth and cell-to-cell spread in a viral context- and cell type-dependent manner. Indeed, in MDCK cells substitution E186K resulted in a strong reduction of U/63 replication rate, a delay in the peak of infection, and an increase in virus titre at the peak of infection (Figure 5-6). However, in the context of O/03, K186E substitution did not change viral growth kinetic (Figure 5-7).

Furthermore, the plaque phenotype of U/63 did not seem to be affected by residue 186 (Figure 5-8). In contrast, K186E substitution significantly reduced O/03 plaque size (Figure 5-9). Furthermore, in E.Derm cells E186K substitution did not improve U/63 viral growth kinetics (Figure 5-10), but significantly reduced O/03 growth abilities (Figure 5-11). These findings are in accordance with the results obtained earlier with the NS reassortant viruses (Figure 5-1). Taken together, these data seem to indicate that in cells that do not possess an efficient antiviral response, like MDCK cells (Seitz et al., 2010), an avian-like virus such as U/63 can tolerate mutations in the NS1 gene that alter important functions of the protein, such as the CPSF30 binding. It is possible that other functions of NS1 or other viral proteins may compensate for the loss of CPSF30 binding of U/63 and allow the virus to maintain fitness in cells permissive for influenza virus. However, the results shown above also highlight that by losing the CPSF30 binding, U/63 lost the capacity to quickly take control of infected cells and rapidly produce new virions, a characteristic that would likely provide a great advantage when jumping hosts. In contrast, for a mammalian-adapted virus, like O/03, viral fitness seems to strongly depend on NS1 and on the maintenance of its evolutionary markers, E186K and C-terminal truncation. The gain of a new NS1 activity (CPSF30 binding) may have disrupted the subtle virus-host equilibrium, leading to attenuation of the virus in interferon competent E.Derm cells. This hypothesis is reinforced by the work of others showing that destabilisation of NS1-CPSF30 complex may arise during the adaptation process of IAV strains to certain hosts (i.e. duck to quail, human to mouse) (Hayman et al., 2006, Kochs et al., 2007a, Twu et al., 2007, Brown et al., 2001, Hossain et al., 2008).

Interestingly, as commented earlier (Chapter 1 section 1.3.1), the surface involved in NS1 ED dimerization is also involved in CPSF30 binding (Hale et al., 2008), and previous work has reported that the disruption of ED-ED interaction by substitution of W187 (direct neighbour of residue 186) resulted in attenuation of a recombinant PR8 virus *in vivo* (Ayllon et al., 2012). The ED dimerization has been shown to play a key role in reinforcing the dsRNA binding of NS1, hence reinforcing NS1's control of the host IFN response. It is possible that restoring the CPSF30 binding in O/03 NS1 has disrupted more important functions that are crucial to support viral replication in IFN competent cells. Additionally, the presence of E186 in O/03 NS1 seemed to have a high fitness cost in equine cells when introduced in the

context of a short NS1 protein, as O/03-K186E replicated poorly in equine cells compared to the O/03-K186E-230 mutant (Figure 5-11). This could indicate that the potential new functions provided by the extension of O/03 NS1 C-terminal tail, e.g. PABPII binding or interaction with PDZ-containing proteins, could partially compensate for fitness cost associated with the reintroduction of CPSF30 binding.

The modification of NS1 C-terminus on its own seemed also to affect EIV infection phenotype, although to a lesser extent than substitution 186. Indeed, the U/63 NS1 truncated mutant virus (U/63-219) displayed a reduced growth rate compared to WT in MDCK cells (Figure 5-6). The growth pattern of U/63-219 resembled the one of U/63-NS-O/03 reassortant virus (Figure 5-1). Furthermore, as for residue 186 substitution, the truncation of NS1 did not seem to affect U/63 plaque phenotype (Figure 5-8), nor improve viral replication in E.Derm cells (Figure 5-10).
In the context of an adapted virus, like O/03, NS1 C-terminal extension alone strongly affected viral replication in interferon competent cells (Figure 5-11), as well as viral spread between neighbouring cells (Figure 5-9). This extension had even more dramatic effects when introduced in association with K186E substitution (Figure 5-7, -9 & -11). Taken together, these data suggest that introduction or suppression of the CPSF30 binding had bigger consequences on viral fitness than a PBM or PABPII binding sites.

When the response of equine cells to O/03 WT and mutant virus infection was assessed, it was observed that only the WT virus was able to replicate to high titres (Figure 5-11) in the presence of high levels of antiviral cytokines (Figure 5-14-A). Moreover, the O/03-230 and O/03-K186E-230 mutant viruses seemed to be able to control better than WT the production of antiviral cytokines upon infection (Figure 5-14-C & -D). These results are in accordance with a previous report from Daly & colleagues, who showed that ponies infected with an EIV from 2003 expressing a truncated version of NS1 (A/equine/Newmarket/5/2003) was inducing larger amounts of type I IFN than an EIV expressing a full length version of NS1 (A/equine/Sussex/1/1989) (Daly et al., 2011, Wattrang et al., 2003). Of note, similar results were found with swine and turkey influenza viruses (Cauthen et al., 2007, Solorzano et al., 2005). Additionally, others reported that decreasing NS1-mediated inhibition of host gene expression correlated with increased innate immune responses after infection (Twu et al., 2007, Hale et al., 2010).

The O/03-WT virus was also able to infect an increasing number of cells and replicate in equine cells primed with universal IFN (Figure 5-16). The ability to replicate in the presence of type I IFN is a significant fitness trait, as it renders an important arm of the host antiviral response ineffective. Moreover, K186 and C-terminal truncation seemed to be detrimental in this process. Interestingly, no Mx1 and ISG15 expression could be detected upon O/03-WT infection in equine cells (Figure 5-13) despite the presence of large amounts of antiviral cytokines secreted (Figure 5-14). This suggests that O/03-WT was able to repress ISG expression, although further work would be needed to determine the mechanism involved. Furthermore, the role of the JAK/STAT pathway in limiting viral replication of O/03-NS1 mutant viruses in E.Derm cells was confirmed with the use of Ruxolitinib prior to viral infection (Figure 5-15).

Finally, when looking at protein production upon infection it was observed that O/03-K186E-230 induced a cellular protein shutdown, particularly at early times post infection (Figure 5-12). This was associated with a premature induction of apoptosis compared to WT (Figure 5-13). This was followed by expression of Mx1 and ISG15 starting at 24 hpi and increasing at 48 hpi (Figure 5-13). The extension of NS1 C-terminus on its own also increased ISG proteins expression compared to WT and resulted in early induction of apoptosis in infected cells (Figure 5-13). However, similarly to WT no protein shutdown was detectable for the O/03-230 mutant virus (Figure 5-12). These results are in accordance with data obtained by others (Golebiewski et al., 2011, Javier and Rice, 2011, Liu et al., 2010) showing that full length NS1 proteins are involved in early induction of apoptosis, likely due to interaction with PDZ-domain containing proteins important in the maintenance of cellular homeostasis (Harris and Lim, 2001, Hung and Sheng, 2002, Kim and Sheng, 2004). It is also possible that the premature induction of apoptosis by O/03-K186E-230 was an indirect consequence of the many cellular-disturbing functions due to CPSF30 binding. Finally, we cannot rule out the possibility that premature apoptosis is due to a lack of control of PKR (Takizawa et al., 1996, Fujimoto et al., 1998, Takizawa et al., 1995, Wada et al., 1995), or a lack of activation of the PI3K/Akt-pathway. Indeed, the presence of viral dsRNA activates PKR and starts a cascade of events eventually leading to protein shutoff (Bergmann et al., 2000, Hatada et al., 1999, Wang et al., 2000) and induction of apoptosis (Takizawa et al., 1996, Van Campen et al., 1989, Gil and Esteban, 2000). NS1 has been shown to inhibit PKR function by several mechanisms (Tan and Katze, 1998, Hatada et al., 1999, Talon et al., 2000a), and delay apoptosis by activation of the PI3K/Akt-pathway (Ehrhardt et al., 2006, Ehrhardt et al., 2007, Hale et al., 2006, Shin et al., 2007c). In the future, it would be interesting to test if E186K substitution and C-terminal truncation modified EIV NS1 blockade

of PKR function, or altered its interaction with the p85β subunit of PI3K, and subsequent activation of PI3K/Akt pathway.

Taken together, these data indicate that the genetic constellation of an equine-adapted virus, like O/03, could already be optimized to function efficiently in equine cells and any disruption in viral-host interaction has a high fitness cost. Interestingly, a similar evolution pattern has been observed for the North American 'classical' swine H1N1 lineage, whose NS segment is of avian origin (related to the 1918 pandemic H1N1 influenza A virus). This virus maintained a full length NS1 protein (230 amino acids) until the mid-1960s, before introducing of a stop codon at position 219 that resulted in the same 11-amino acid C-terminal truncation of than EIV NS1. These changes have subsequently been retained in the 'classical' swine H1N1 lineage until the present day (Hale et al., 2010). Thus, the loss of CPSF30 binding and C-terminal truncation of NS1 seem to be a common evolutionary trait between swine and equine Influenza A viruses. Interestingly, human influenza A viruses seem somewhat different, and preferentially select for NS1 proteins binding CPSF30. For example, the NS1 protein of H5N1 viruses presented a defect in inhibiting general gene expression when the transmission occurred from birds to humans in 1997, however the viruses isolated since 1998 have gained this function (Twu et al., 2007, Clark et al., 2017). It would be interesting to compare the advantages and downsides of losing CPSF30 binding and variations of the C-terminal in different mammalian hosts, such as horses, pigs, dogs and humans.

Chapter 6

Effect of NS1 residue 186 and C-terminus on virus-host interaction

6. Effect of NS1 residue 186 and C-terminus on virus-host interaction

6.1. Introduction

The interface between an infectious agent and its host represents an ultimate battleground for survival, as the virus must secure a niche for replication, while the host must limit the pathogen's advance. To mount a productive antiviral response against IAVs, multiple intracellular signalling pathways must be activated, and the virus can exploit these pathways to support its own replication (Ludwig et al., 2006, Ludwig et al., 2003, Ehrhardt et al., 2007, Kochs et al., 2007). This is notably the task of the NS1 protein, an accessory factor important for the blockade of the host innate immune response, which targets the IFN system at different steps. NS1 can limit induction of IFN by blocking the activation of transcription factors such as IRF3 (Talon et al., 2000), NF-κB (Wang et al., 2000) and AP-1 (Ludwig et al., 2002), or by interfering with the RIG-I signalling pathway (Pichlmair et al., 2006, Mibayashi et al., 2007) notably through the binding to TRIM25 (Gack et al., 2009) and RIPLET (Rajsbaum et al., 2012). NS1 is also important for taking over the cellular machinery, shutting down the host's own protein production, including those of type I IFN, while promoting viral-only gene expression. This is notably achieved by blocking splicing of pre-mRNA and export of poly-adenylated mRNA molecules out of the nucleus through the binding to CPSF30 (Nemeroff et al., 1998, Das et al., 2008) and PABPII (Chen et al., 1999). The dsRNA binding ability of NS1 could also play a role in pre-transcriptional inhibition of the interferon pathway by sequestering PAMPs away from host viral sensors (Pichlmair et al., 2006). More recently, NS1 RNA-binding activity has been involved in the control of gene expression by preventing the loading of the transcriptional machinery to the host DNA (Anastasina et al., 2016), and by suppressing RNA interference and RNA silencing activities (Li et al., 2016b). These NS1 pro-viral activities will ultimately disrupt host cell homeostasis and promote cell death. Interestingly, NS1 has also developed mechanisms to interfere with the PI3K signalling network (Hale et al., 2008b, Hale et al., 2006) in order to increase infected cell viability and maintain the availability of cellular resources. However, many of these NS1 functions are not conserved between IAV strains (de Vries et al., 2009), and this will likely impact on virus-host interaction and on virus replication efficiency.

In the previous chapters, two evolutionary markers of NS1's function were identified, e.g. NS1 E186K substitution and C-terminal truncation. These markers were further confirmed as important for EIV infection phenotype, and their impact seemed to be viral context- and cell type-dependent. Indeed, viral fitness of a mammalian-adapted EIV (O/03) was strongly dependent on the maintenance of NS1 evolutionary markers, as the

reintroduction of the CPSF30 binding via K186E substitution led to viral attenuation. A similar observation was made for the extension of NS1 C-terminus (Chapter 5). These changes resulted in a reduction of infectivity, in the incapacity to replicate in the presence of antiviral cytokines, in the inability to block the establishment of an anti-viral state and in premature induction of apoptosis. In contrast, for an emergent virus like U/63, the NS1 C-terminal truncation did not strongly modify viral infection phenotype, however, the loss of CPSF30 binding (via E186K substitution) was associated with a reduction in replication rate early post infection. Although the effect of NS1 residue 186 and C-terminus clearly affected virus-host interaction in a viral-context dependent manner, the mechanisms involved remained elusive. A better understanding of the global gene changes following wild-type and mutant virus infection could help identify cellular pathways involved.

High-throughput RNA sequencing (RNA-seq) technology is a powerful way to analyse the transcriptome of infected cells with high efficacy and accuracy (Josset et al., 2014, Park et al., 2015, Wang et al., 2015). Limited work associating reverse genetic techniques and gene expression profiling in infected cells is available, but Geiss et al. reported that NS1 C-terminus was important for controlling host transcriptome and the initial cellular antiviral response. They notably demonstrated that NS1 deletion mutants (full deletion or important C-terminal deletions) were unable to alter host gene expression profile, and this was associated with a rapid clearance of the virus *in vitro* (Geiss et al., 2002).

In this chapter, a global transcriptomic profiling to equine cells infected with O/03 and U/63 WT and isogenic mutants for codon 186 and/or codon 220 was applied in order to broadly characterize the impact of NS1 evolutionary markers on host response to viral infection.

Aims: The aims of this chapter were to compare the early genes expression pattern of equine cells infected with two evolutionary distinct EIVs, and to evaluate the impact of NS1 evolutionary markers on virus-host interaction.

6.2. Results

6.2.1. Preparation of samples for RNA sequencing

E.Derm cells were infected with O/03 (Figure 6-1A) and U/63 (Figure 6-1B) WT and mutant viruses or mock infected them for 8h (end of eclipse phase) (Figure 5-7) in order to avoid any bias due to different replication efficiency between viruses. The MOI was adjusted for each independent infection to ensure that a similar number of cells were infected in all cells. The level of infection in each sample was then analysed by evaluating the percentage of live cells expressing the viral nucleoprotein by flow cytometry, as described in Chapter 3, section 3.5.7.

In O/03-infected cells (Figure 6-1A), the percentage of cells expressing the viral NP among live cells was between 23.1% to 31.6%, and the mean cellular viability ranged from 77.51% to 88.97%. In U/63-infected cells (Figure 6-1B), the percentage of live cells expressing the viral NP ranged between 16.8% and 29.0%, with a cellular viability ranging between 77.59% and 88.97%. Importantly, U/63 double-mutant virus had significantly fewer infected cells compared to the two other mutant viruses, thus it was not possible to compare the transcriptome of single mutant-infected cells with the one of double mutant virus-infected cells.
Of note, the majority of cells (approximately 70%) for both O/03 and U/63 infected cells whose transcriptome was analysed were not infected.

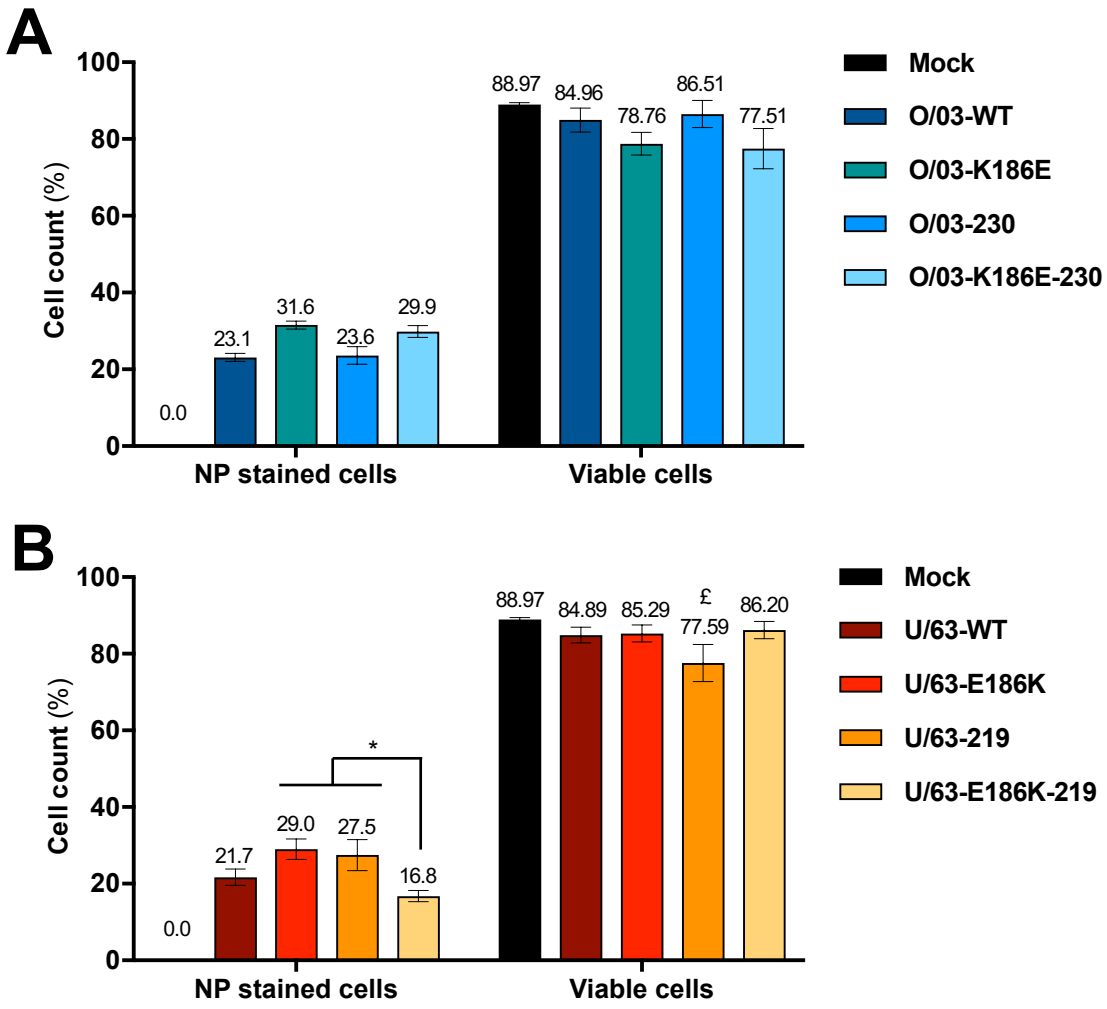


Figure 6-1: Level of infection and cell viability

E.Derm cells were infected with (**A**) O/03- and (**B**) U/63-WT and mutant viruses or mock infected for 8 h. The percentage of cells expressing the viral nucleoprotein among live cells and total cell viability was determined by flow cytometry. Bars correspond to mean of three independent experiments and error bars represent SEM. A Two-way ANOVA with Bonferonni post hoc test was used to determined significance. *, $p<0.05$ for U/63-E186K-219 versus indicated cells. £, $p<0.05$ for U/63-219 versus mock-infected cells. No statistical difference was observed for O/03-infected cells.

Total RNA was then extracted for each sample, and the RNA quality was analysed using a bioanalyzer, as described in Chapter 3, section 3.9. All cells had an RNA Integrity Number (RIN) above 9.5 on a scale from 1 to 10 (Figure 6-2).

The RNA samples were then sequenced and analysed as described in Chapter 3, section 3.9.

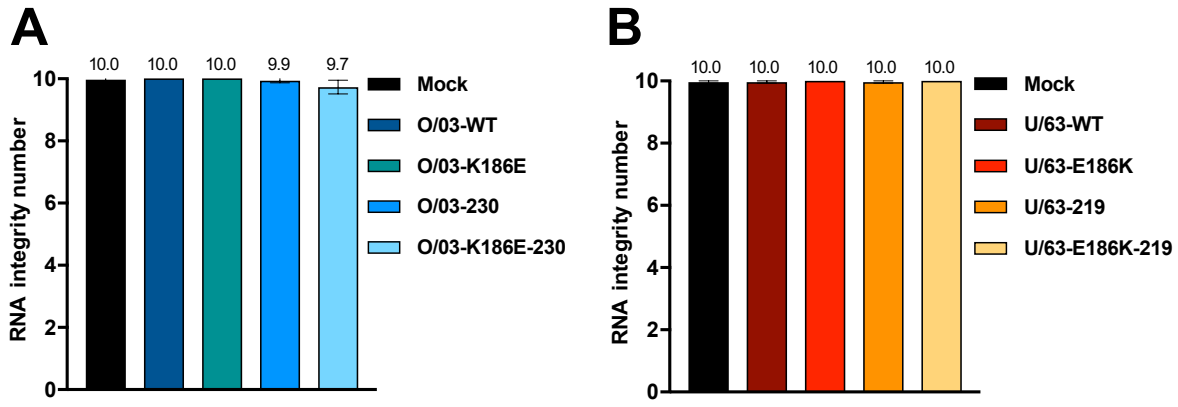


Figure 6-2: RNA Integrity Number (RIN)

The average RIN of mock-infected or (**A**) O/03- and (**B**) U/63-WT and mutant viruses infected cells. Bars correspond to mean of three independent experiments and error bars represent SEM. No significant difference was found.

6.2.2. Transcriptome of O/03 wild-type and mutant virus-infected cells

6.2.2.1. Degree of dissimilarity between samples

To evaluate the degree of dissimilarity between mock and virus-infected cells and among biological replicates, a multi-dimensional scaling (MDS) plot was generated (Figure 6-3). This plot is used to visualize Euclidian distances between each transcriptomic profile in two dimensions (Josset et al., 2014). In this plot, each individual sample is represented by a dot, and cells with similar transcriptomic profiles (small Euclidian distance) will be found close to each other, while increasing distances between dots reveal increasing transcriptomic dissimilarities. Overall, all biological replicates for each condition were closely distributed. In addition, O/03-WT-, O/03-K186E-, and O/03-K186E-230-infected cells were closely clustered, while O/03-230-infected cells were clustering with mock-infected cells.

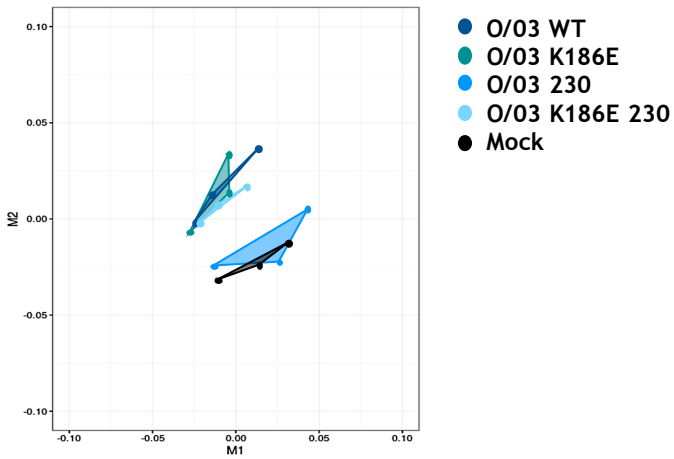


Figure 6-3: Multidimensional scaling plot for O/03-WT- and mutant-infected samples.

Variation between cells and biological replicates was analysed using a multidimensional scaling (MDS) plot.

6.2.2.2. Distribution of sequencing reads

The total number of reads as well as the number of reads aligning with the horse genome (Equus caballus 2, GCA_000002305.1) or the viral genome (A/equine/Ohio/1/2003) were analysed in each sample (Figure 6-4). A mean of 41.2 million reads were obtained per sample, with 36.6 million horse genome reads, and 3.5 million viral genome reads.

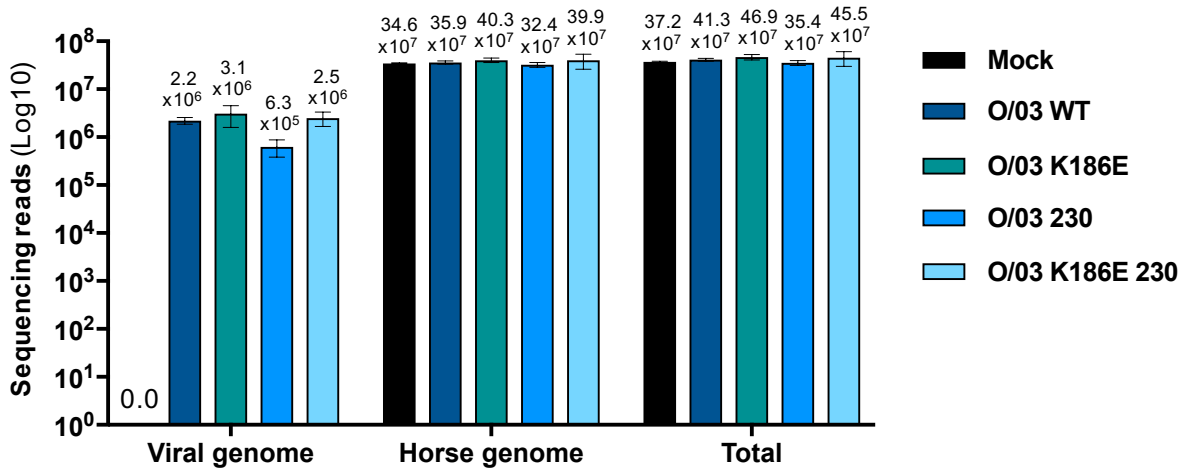


Figure 6-4: Average number of reads in O/03-WT- and mutant-infected samples.

The mean of total number of reads was determined for each condition, as well as the mean of reads matching the viral genome or the horse genome. The exact numbers of reads for each condition is indicated. Bars correspond to mean of three independent experiments and error bars represent SEM.

6.2.2.3. Viral reads

To determine if the mutations introduced in NS1 affected viral gene expression, the transcript per kilobase million (TPM) per genomic segment was compared between conditions. As seen in (Figure 6-5), no significant differences was observed.

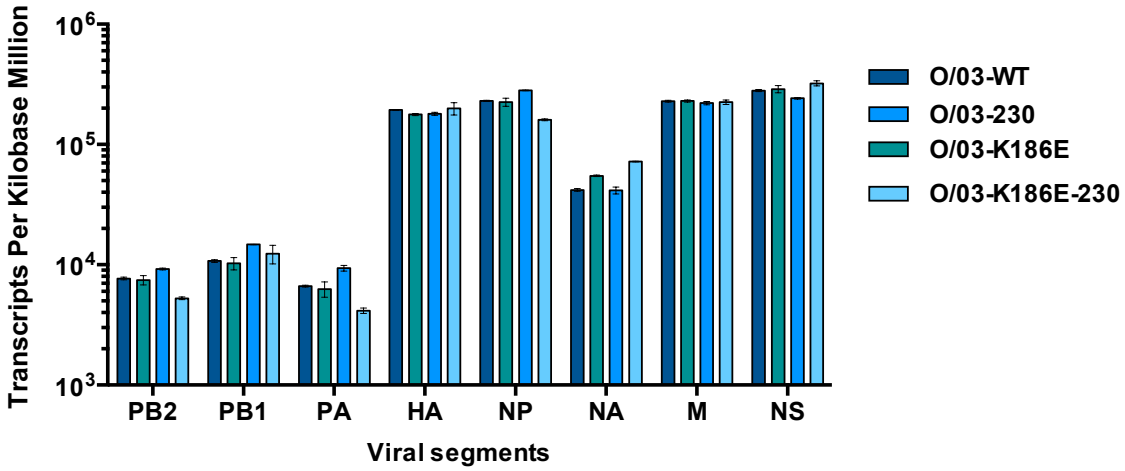


Figure 6-5: Sequencing reads per viral genomic segment in O/03-WT- and mutant-infected cells.

To compare the proportion of reads that mapped to specific viral genomic segment in each condition transcripts per kilobase million were calculated by normalizing sequencing reads for the respective gene length and sequencing depth. Bars correspond to mean of three independent experiments and error bars represent SEM. No significant difference was found using by Two-way ANOVA with Bonferonni post hoc test.

6.2.2.4. Differential gene expression between O/03 virus- and mock-infected cells

The total number of differentially expressed genes (DEGs) in virus-infected cells compared to mock-infected cells (Figure 6-6) was then evaluated. A total of 429 DEGs compared to mock were observed in O/03-WT-infected cells, with 248 genes up-regulated and 181 down-regulated. In O/03-K186E-infected cells 193 genes were differentially expressed compared to mock, with 65 up-regulated and 128 down-regulated. In O/03-230-infected cells, a total of 241 genes were differentially expressed compared to mock, with 179 up-regulated and 62 down-regulated. Finally, in O/03-K186E-230-infected cells, 158 genes were differentially expressed compared to mock, with 101 up-regulated genes and 57 down-regulated genes. The entire list of DEGs for each condition is available in Appendices 35 to 45.

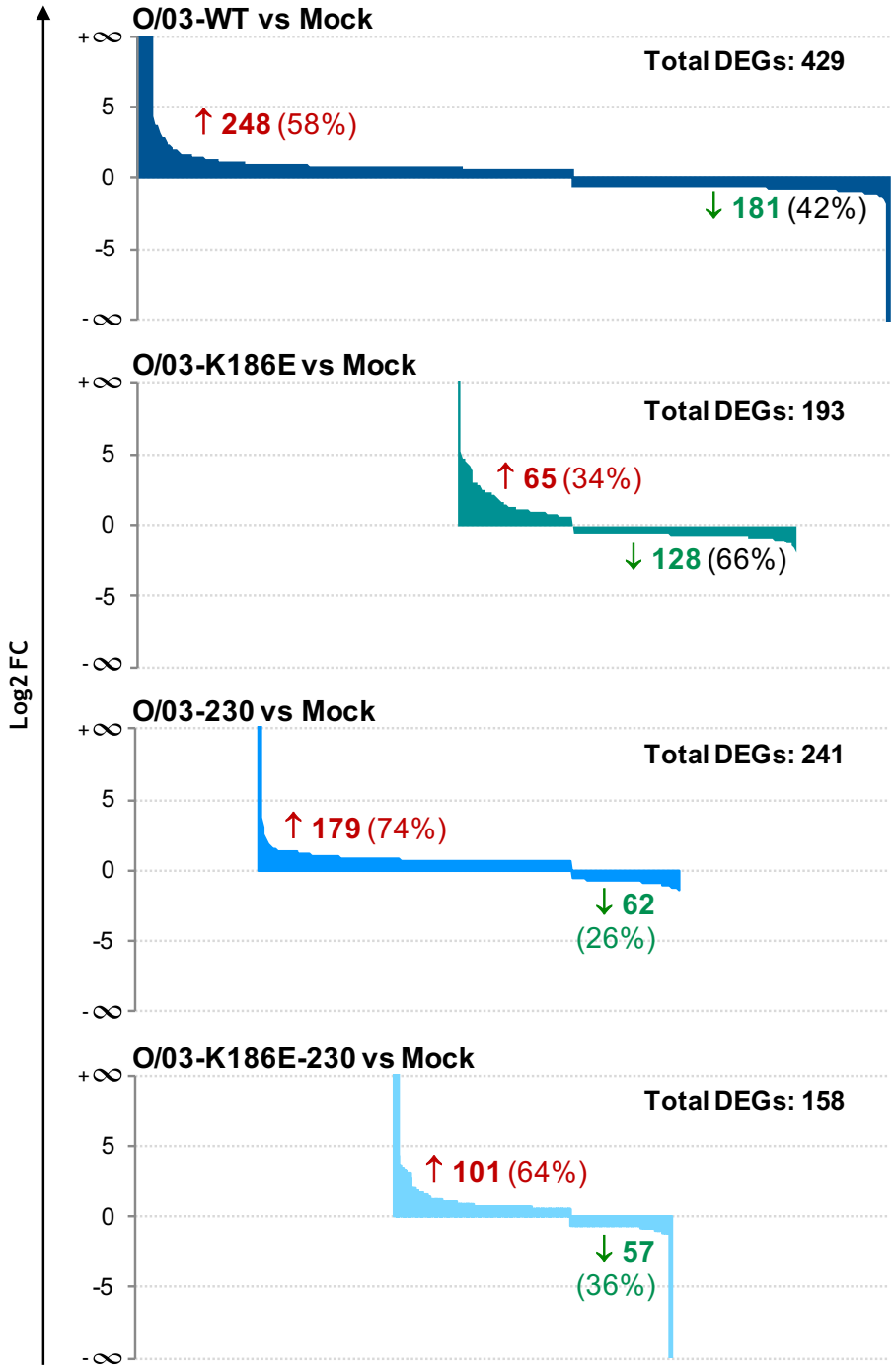


Figure 6-6: Plots representing DEGs compared to mock in O/03-WT- and mutant-infected cells.
Each plot displays the range of host genes found significantly differentially expressed compared to mock-infected cells (Log2FC≥0.58 and p value<0.05) in cells infected with each virus. Total number of DEGs are indicated in black, and up-regulated genes and down-regulated genes are indicated in red and green (with total number in bold and percentage over total DEGs in brackets), respectively.

The degree of overlap between gene sets was then analysed (Figure 6-7). It was observed that the O/03-WT virus alone modified the expression of the largest number of cellular genes (171 up-regulated and 113 down-regulated), whereas the two K186E-containing mutant viruses alone regulated the fewest number of genes (22 up-regulated and 59 downregulated for O/03-K186E and 25 up-regulated and 22 down-regulated for O/03-K186E-230). The O/03-230 virus regulated an intermediate number of cellular genes, with 127 up-regulated and 44 down-regulated.

Two genes were regulated in opposite directions by O/03-WT and the single mutant viruses. The Heat shock 70 kDa protein 6 (HSPA6) was up-regulated in O/03-WT-infected cells but down-regulated in O/03-K186E (Log2FC = 1.50962 and -1.31121, respectively). This gene is the 6th member of the Heat Shock Protein family A (Hsp70), whose induction is a direct consequence of HSP90 inhibition (Kuballa et al., 2015), a chaperone that plays a key role in many important cellular processes, including DNA repair, development, immune response (Schopf et al., 2017). The Urocortin-2 (UCN2) was down-regulated in O/03-WT and up-regulated in O/03-230-infected cells (Log2FC = -1.089180 and 0.835742, respectively). This gene encodes for a peptide previously shown to induce the expression of IL-8 and promote cell survival and proliferation (Lee et al., 2014, Shi et al., 2006).

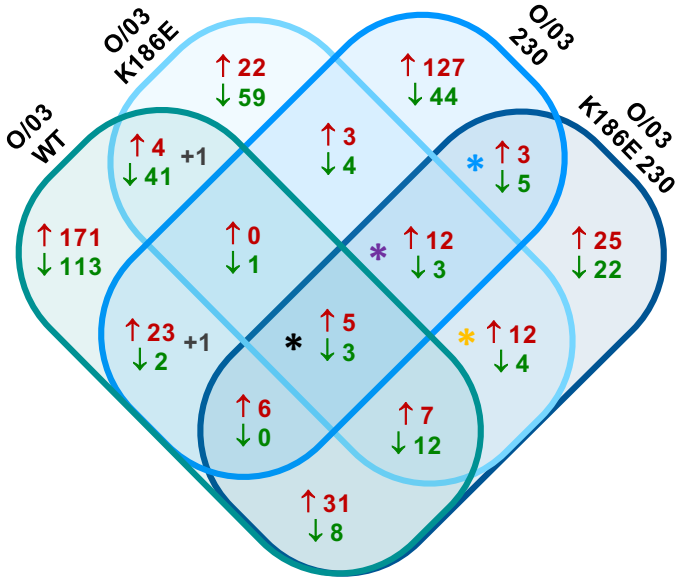


Figure 6-7: Venn diagram representing DEGs compared to mock in common between O/03-infected samples.

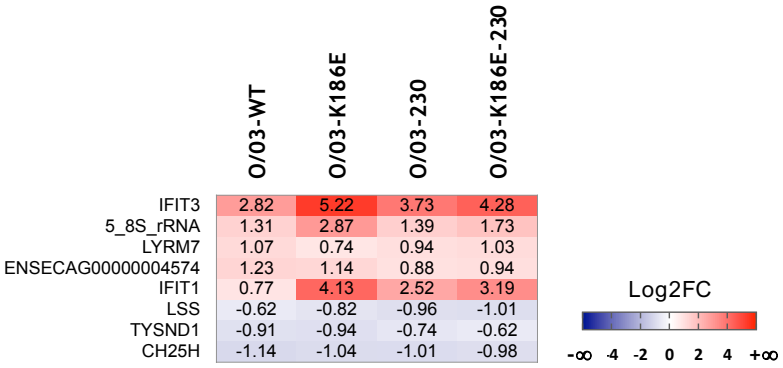
DEGs compared to mock in infected cells with a Log2FC≥0.58 and p value<0.05 were used this study. Up-regulated and down-regulated genes are indicated in red and green, respectively, while a black number indicates a gene differentially expressed in the opposite direction in two different conditions. Asterisks indicate four primary groups of DEGs: black asterisk (*) indicates DEGs of primary group O1, purple asterisk (*) indicates DEGs of primary group O2, yellow asterisk (*) is for DEGs of primary group O3, and blue asterisk (*) are DEGs of primary group O4.

From the gene sets overlap analysis (Figure 6-7), four primary groups were identified: **(O1)** genes whose expression was similar between WT and mutant virus infected cells (black asterisk); **(O2)** genes whose expression was similar between cells infected with mutant viruses but absent in WT-infected cells (purple asterisk); **(O3)** genes whose expression was differentially regulated compared to mock in K186E-containing mutant virus-infected cells (yellow asterisks); and **(O4)** genes whose expression was differentially regulated compared to mock in 230-containing mutant virus-infected cells (blue asterisks).

In group O1 (Figure 6-8), five genes were up-regulated and three down-regulated. Among up-regulated genes, two IFN-induced genes involved in the cellular response to viral infection (interferon-induced protein with tetra-tricopeptide repeats 3 and 1, IFIT3 and IFIT1, respectively) were identified (Pinto et al., 2015, Moll et al., 2011). Both of these genes were strongly up-regulated in mutant-infected cells, particularly with O/03-K186E, and less so in O/03-WT-infected cells.

The function of the three other up-regulated genes remains unknown: LYRM7 (LYR motif containing 7), a gene coding for a component of the mitochondrion; 5_8S_rRNA, a novel ribosomal RNA; and ENSECAG00000004574, an unannotated gene.

Among the three down-regulated genes shared between all conditions was the LSS gene (lanosterol synthase), encoding a key intermediate in cholesterol biosynthesis (Chen et al., 2015); TYSND1 (trypsin domain containing 1), a key regulator of the peroxisomal β-oxidation pathway (Okumoto et al., 2011); and CH25H (cholesterol 25-hydrolase), a known ISG involved in fatty acid and cholesterol biosynthetic processes (Gold et al., 2014, Chen et al., 2014, Serquina et al., 2017). These genes were regulated in a similar fashion between samples.



Gene	Key words/functions
IFIT3	Defense response to virus and cellular response to type I IFN, negative regulation of cell proliferation
5_8S_rRNA	Novel ribosomal RNA, unknown function
LYRM7	Component of mitochondrion, unknown function
ENSECAG00000004574	Uncharacterized protein
IFIT1	Defense response to virus and cellular response to type I IFN. Intracellular transport of viral protein in host cell, negative regulation of protein binding and helicase activity
LSS	Steroid biosynthetic process, cholesterol biosynthetic process, regulation of protein stability
TYSND1	Regulation of fatty acid beta-oxidation
CH25H	Fatty acid biosynthesis process, cholesterol metabolic process and oxidation-reduction process

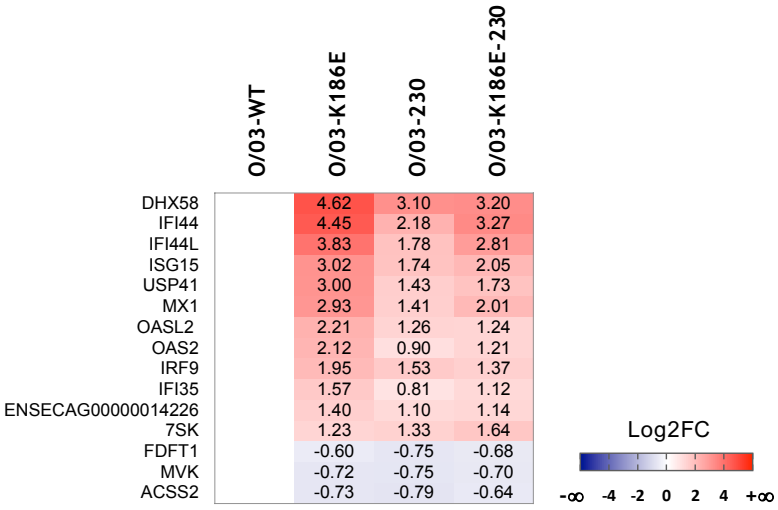
Figure 6-8: Heatmap highlighting DEGs compared to mock in O/03-WT and mutant virus-infected cells.

Genes whose expression is statistically over-represented (up or down-regulated with Log2FC≥0.58 and p value<0.05) in all virus-infected cells compared to mock-infected ones.

In group O2 (Figure 6-9), twelve genes were up-regulated and the majority were associated with the type I IFN pathway and the regulation of host defence against viral infection. Among these genes were DExH-box helicase 58 (DHX58), also known LGP2, interferon-induced protein 4 (IFI44), interferon-induced protein 44 like (IFI44L), ISG15, ubiquitin specific peptidase (USP41), Mx1, OAS-like protein 2 (OASL2), OAS2, IRF9, and interferon-induced 35 kDA (IFI35) (Malur et al., 2012, Hallen et al., 2007, Moal et al., 2013, Haller et al., 2007, Li et al., 2016a, Das et al., 2014, Qashqari et al., 2013, Fan et al., 2017). These genes were strongly expressed in O/03-K186E-infected cells.

Two other genes were upregulated: the 7SK gene, a non-coding RNA that positively regulates the transcription elongation factor, P-TEFb, which plays an essential role in the regulation of RNA polymerase II transcription in eukaryotes; and an uncharacterized gene (ENSECAG00000014226).

Furthermore, three genes were down-regulated: FDFT1 (farnesyl-diphosphate farnesyltransferase 1) (Serquina et al., 2017), MVK (mevalonate kinase (Fenini et al., 2017)), and ACSS2 (Acyl-CoA synthetase short-chain family member 2, (Xu et al., 2017)), all involved in lipid metabolism and regulation of inflammatory processes.



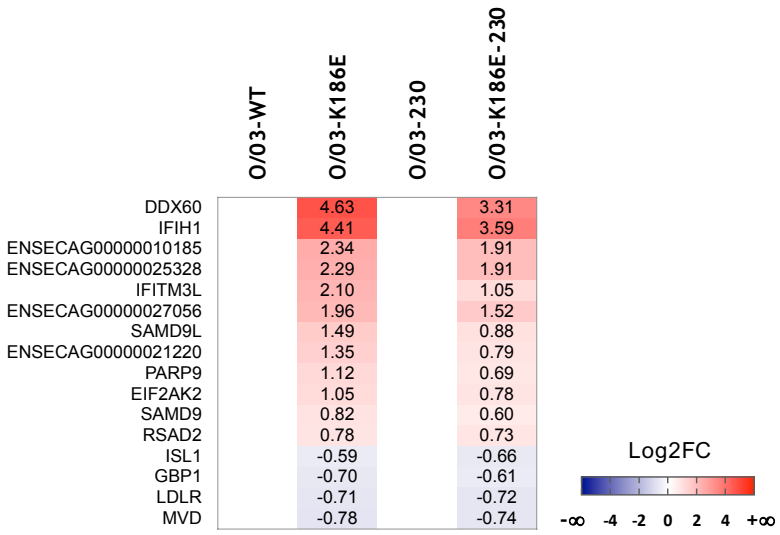
Gene	Key words/functions
DHX58	Negative and positive regulation of MDA-5 and RIG-I signaling pathways
IFI44	Defense response to virus
IFI44L	Defense response to virus
ISG15	Defense response to virus, regulation of IFN-gamma production, modification-dependent protein catabolic process, negative regulation of protein ubiquitination
USP41	Proteasome-mediated ubiquitin-dependent protein catabolic process
MX1	Defense response to virus, negative regulation of viral genome replication
OASL2	Defense response to virus, purine nucleotide biosynthetic process
OAS2	Defense response to virus, RNA catabolic process, protein glycosylation
IRF9	Transcription factor activity, type I IFN signaling pathway
IFI35	Type I IFN signaling pathway
ENSECAG00000014226	Uncharacterized protein
7SK	Novel non-coding RNA that regulates the general transcription factor P-TEFb
FDFT1	Lipid biosynthetic process
MVK	Lipid metabolic process, negative regulation of inflammatory response
ACSS2	Metabolic process, lipid biosynthetic process, acetyl-CoA biosynthetic process

Figure 6-9: Heatmap highlighting DEGs compared to mock in O/03 mutant virus-infected cells.

Genes whose expression is statistically over-represented (up or down-regulated with Log2FC≥0.58 and p value<0.05) in mutant virus-infected cells compared to mock-infected condition.

In group O3 (Figure 6-10), twelve genes were up-regulated, with higher Log2FC values in single mutant- than in double mutant virus-infected cells. These up-regulated genes were mainly involved in the type I IFN pathway, such RIG-I (also known as DEXD/H box helicase 60, DDX60), MDA-5 (also known as interferon-induced with helicase C domain, IFIH1), interferon-induced transmembrane protein 3-like (IFITM3L), PKR (or eukaryotic translation initiation factor 2 alpha kinase 2, EIF2AK2), viperin (also known as radical S-adenosyl methionine domain containing 2, RSAD2), guanylate binding protein 3 (GBP3) and poly-ADP-ribose polymerase family member 9 (PARP9), and sterile alpha motif domain containing 9 and 9-like (SAMD9 and SAMD9L) (Zhang et al., 2015, Wang et al., 2008, Sato et al., 2015, Kato et al., 2006, Haller et al., 2015, Hale et al., 2008a, Liu and McFadden, 2015). Additionally, three up-regulated genes were not annotated: ENSECAG00000010185, ENSECAG00000025328, ENSECAG00000027056.

In this group, four genes were down-regulated: guanylate-binding protein (GBP1), a known ISG (Niu et al., 2016); low-density lipoprotein receptor (LDLR), described as an entry port for different viruses (Chen et al., 2014, Finkelshtein et al., 2013); mevalonate diphosphate decarboxylase (MVD), previously shown to favour Dengue virus replication (Rothwell et al., 2009); and insulin gene enhancer protein 1 (ISL1) involved in RNA pol II activation (Whitney et al., 2015).



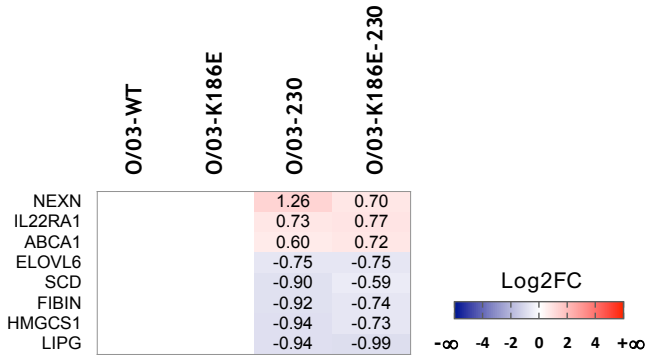
Gene	Key words/functions
DDX60	Defense response to virus, positive regulation of MDA-5 and RIG-I signaling pathways
IFIH1	Defense response to virus, protein sumoylation, positive regulation of type I IFN production, MDA-5 signaling pathway, regulation of apoptotic process
ENSECAG00000010185	Novel pseudogene, unknown function
ENSECAG00000025328	Mi-RNA, unknown function
IFITM3L	IFN-induced transmembrane dispanin subfamily A member
ENSECAG00000027056	Mi-RNA, unknown function
SAMD9L	Endosomal vesicle fusion
GBP3	Response to type I IFN, GTPase activity, unknown function
PARP9	Double-strand break repair, regulation of response to IFNγ
EIF2AK2	Response to type I IFN, positive regulation of cytokine and chemokine production, negative regulation of translation and viral genome replication, regulation of hematopoietic stem cell proliferation and differentiation
SAMD9	Endosomal vesicle fusion
RSAD2	Defense response to virus, positive regulation of TLR7 and TLR9 signaling pathways, CD4-positive, α/β-T cell activation and differentiation, positive regulation of T-helper 2 cell cytokine production
ISL1	RNA pol II activating transcription factor binding, RNA pol II transcription co-activator activity, enhancer sequence-specific DNA binding
GBP1	Response to type I IFN, nucleotide binding, GTPase activity
LDLR	Amyloid-beta binding, protease binding, low-density lipoprotein receptor activity, low-density particle binding
MVD	Hsp70 protein binding, ATP binding

Figure 6-10: Heatmap highlighting DEGs compared to mock in O/03 K186E-containing mutant virus-infected cells.

Genes whose expression is statistically over-represented (up or down-regulated with Log2FC≥0.58 and p value<0.05) in O/03 NS1-K186E-containing mutant virus-infected cells compared to mock-infected condition.

In group O4 (Figure 6-11), three genes were up-regulated and five downregulated. Among up-regulated genes was found the gene encoding for the Nexilin F-actin binding protein (NEXN), a known ISG in primary B cells (Khsheibun et al., 2014); the interleukin 22 receptor subunit alpha 1 gene (IL22RA1), involved in lung repair post influenza infection (Noyama et al., 2017, McAleer and Kolls, 2014); and the ATP-binding cassette subfamily A member 1 gene (ABCA1), which encodes for a protein that modulates the bioavailability of cholesterol in the lipid raft and favours viral entry (Pirillo et al., 2015).

Among the five down-regulated genes, was the gene coding for Fibin, a protein of the endoplasmic reticulum and Golgi involved in wnt signalling and organogenesis in zebrafish (Wakahara et al., 2007). The four other genes were involved in lipid homeostasis, such as the elongation of long-chain fatty acids family member 6 gene (ELOVL6), which encodes an enzyme that catalyses saturated fatty acids elongation and has been associated with lung homeostasis regulation (Sunaga et al., 2013); the stearoyl-CoA desaturase gene (SCD), encoding a protein involved in the biosynthesis of unsaturated fatty acids (Lyn et al., 2014); the 3-hydroxy-3-methylglutaryl-CoA synthase 1 gene (HMGCS1), which is involved in the mevalonate/cholesterol pathway (Serquina et al., 2017); and a gene encoding for the endothelial lipase G (LIPG), an enzyme determining plasma lipoprotein cholesterol (Otera et al., 2009).



Gene	Key words/functions
NEXN	Regulation of cell migration and cytoskeleton organization
IL22RA1	Cytokine receptor activity, defense response to Gram-negative bacterium
ABCA1	ATPase activity coupled to transmembrane movement of substances
ELOVL6	Fatty acid elongase activity, acyl groups transferase activity
SCD	Oxidoreductase activity
FIBIN	Protein associated with the endoplasmic reticulum and Golgi apparatus, unknown function
HMGCS1	Isoprenoid biosynthetic process
LIPG	Cholesterol homeostasis, phospholipid homeostasis, regulation of lipoprotein metabolic process

Figure 6-11: Heatmap highlighting DEGs compared to mock in O/03 230-containing mutant virus-infected cells.

Genes whose expression is statistically over-represented (up or down-regulated with Log2FC≥0.58 and p value<0.05) in O/03 NS1-230-containing mutant virus-infected cells compared to mock-infected condition.

6.2.2.5. Top canonical signalling pathways modified by O/03 viruses

To identify possible biological interactions of DEGs and important functional networks, the sets of genes with altered expression in O/03-WT and mutant virus infected cells were imported into the Ingenuity pathway analysis (IPA) tool. As seen in (Figure 6-12), the O/03-WT virus activated a larger number of pathways compared to mutant viruses. Infection with this virus was associated with up-regulation of genes involved in cell survival, such as the PI3K/AKT pathway, and down-regulation of genes associated with cell cycle arrest and cell death (i.e. apoptosis signalling, telomerase signalling, cell cycle G1/S checkpoint regulation, p53 signalling, role of BRCA1 in DNA damage response, cell cycle G2/M checkpoint regulation). Some inflammatory pathways were also downregulated in these cells (i.e. IL-17A signalling in airway cells, IL-6 signalling, LXR/RXR activation or TNFR1 signalling), while other were activated (i.e. LPS/IL-1 mediated inhibition of RXR function, FcεRI signalling, IL-8 signalling) (Al-Garawi et al., 2012).

In addition, in O/03 mutant virus-infected cells several key innate immune receptor pathways involved in virus sensing were activated, such as interferon signalling, activation of IRF by cytosolic pattern recognition receptors, RIG-I-like receptors and antiviral innate immunity, or pattern recognition receptors of bacteria and viruses.

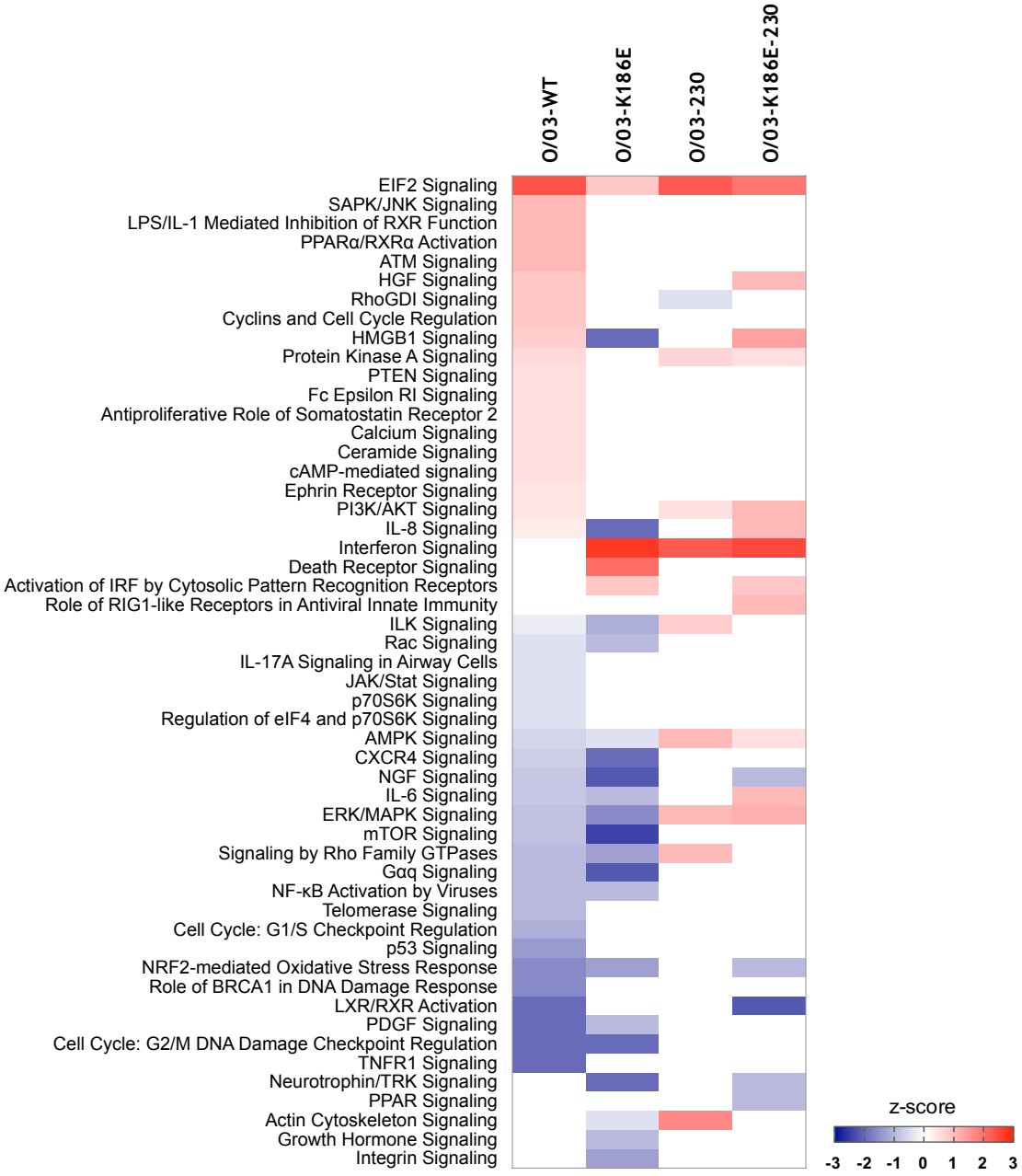


Figure 6-12: Top canonical signalling pathways activated by O/03-WT and mutant viruses.

The Ingenuity Pathway Analysis (IPA) tool was used to generate a list of significant canonical pathways activated in virus-infected cells compared to mock-infected cells based on the lists of transcripts found as significantly differentially expressed compared to mock-infected cells.

The U/63 WT and mutant-infected samples were analysed in a similar fashion to O/03-infected samples.

6.2.3. Transcriptome of U/63 wild-type and mutant virus-infected cells

6.2.3.1. Degree of similarity between samples

The degree of dissimilarity between conditions and biological replicates were analysed with a MDS plot (Figure 6-13). All biological replicates were distributed relatively closely to each other, although not as closely as observed for O/03-infected cells (Figure 6-3). In addition, U/63-WT and U/63-E186K-infected cells were closely clustered, while U/63-219-infected cells were clustering with mock-infected cells. In addition, U/63-E186K-219-infected cells were equally distributed between U/63-WT- and mock-infected cells.

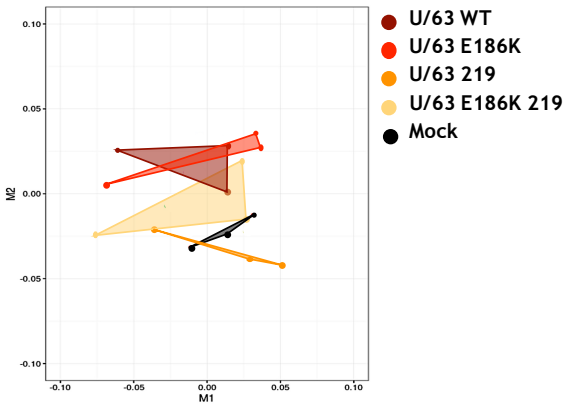


Figure 6-13: Multidimensional scaling plot for U/63-WT- and mutant-infected samples.
Variation between cells and biological replicates was analysed using a multidimensional scaling (MDS) plot.

6.2.3.2. Distribution of sequencing reads

The total number of reads as well as the number of reads aligning to the horse or viral genome were determined for all conditions. As seen in (Figure 6-14), all samples had an average of 35.2 million reads in total, with 32.3 million horse genome reads, and 0.4 million viral genome reads.

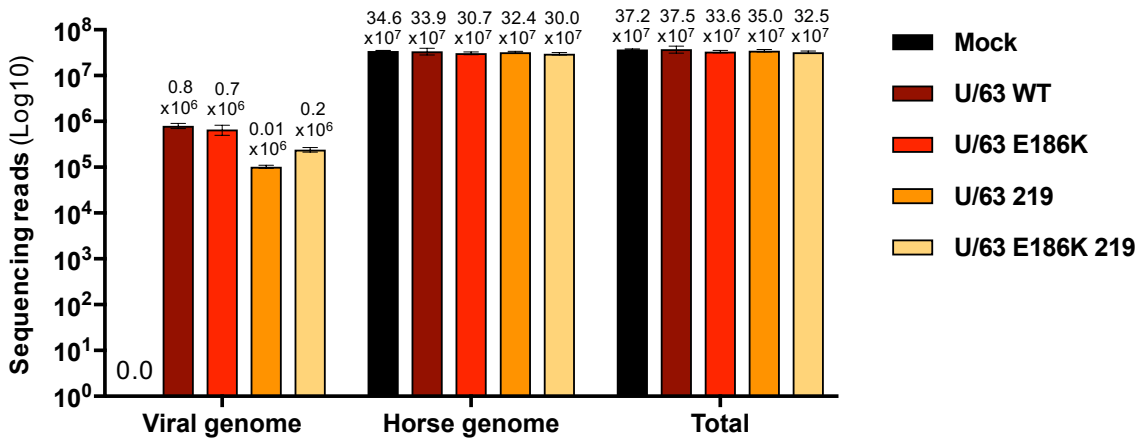


Figure 6-14: Average number of reads in U/63-WT- and mutant-infected samples.

The mean of total number of reads was determined for each condition, as well as the mean of reads matching the viral genome or the horse genome. The exact numbers of reads for each condition is indicated. Bars correspond to mean of three independent experiments and error bars represent SEM.

6.2.3.3. Viral reads

When the viral reads per genomic segments were compared between in all conditions, it was observed that expression of U/63-WT PB2 gene was significantly higher than that of all mutant viruses. In addition, U/63-WT NA gene expression was significantly increased compared to E186K-containing mutants (Figure 6-15).

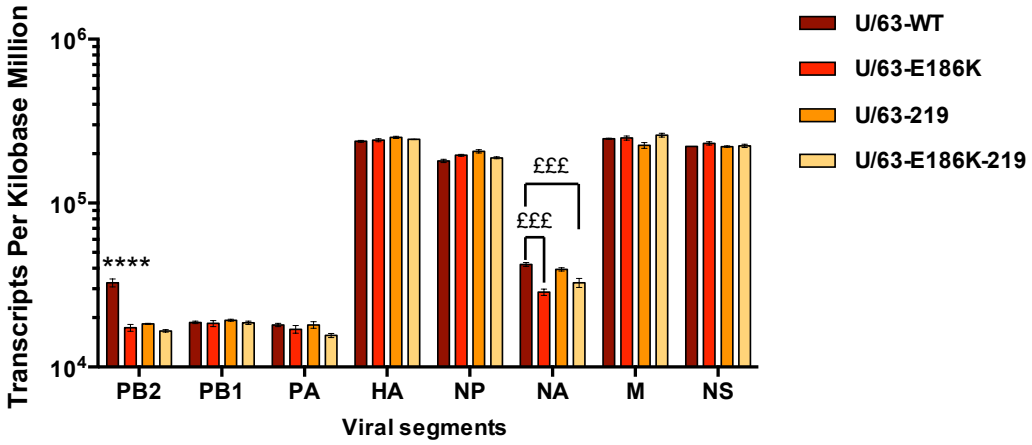


Figure 6-15: Sequencing reads per viral genomic segment in U/63-WT- and mutant-infected samples.

To compare the proportion of reads that mapped to specific viral genomic segment in each condition transcripts per kilobase million were calculated by normalizing sequencing reads for the respective gene length and sequencing depth. Bars correspond to mean of three independent experiments and error bars represent SEM. Significance was calculated by Two-way ANOVA with Bonferonni post hoc test. ****, $p<0.0001$ for indicated U/63-WT versus all other conditions. £££, $p<0.001$ and £, $p<0.05$ for U/63-WT versus E186K-containing mutant virus.

6.2.3.4. Differential gene expression between U/63 virus and mock-infected cells

The total number of DEGs in U/63-WT and mutant virus-infected cells was then evaluated (Figure 6-16). In total, U/63-WT-infected cells showed 398 DEGs compared to mock, among which 140 were up-regulated and 258 were down-regulated.
In U/63-E186K-infected cells, 205 genes were differentially expressed compared to mock with 146 up-regulated and 59 down-regulated. In U/63-219-infected cells there was 216 DEGs compared to mock, with 117 up- and 99 down-regulated genes. Interestingly, in U/63-E186K-219-infected cells, a total of 1126 genes were differentially expressed compared to mock, with 1060 up-regulated and 66 down-regulated.

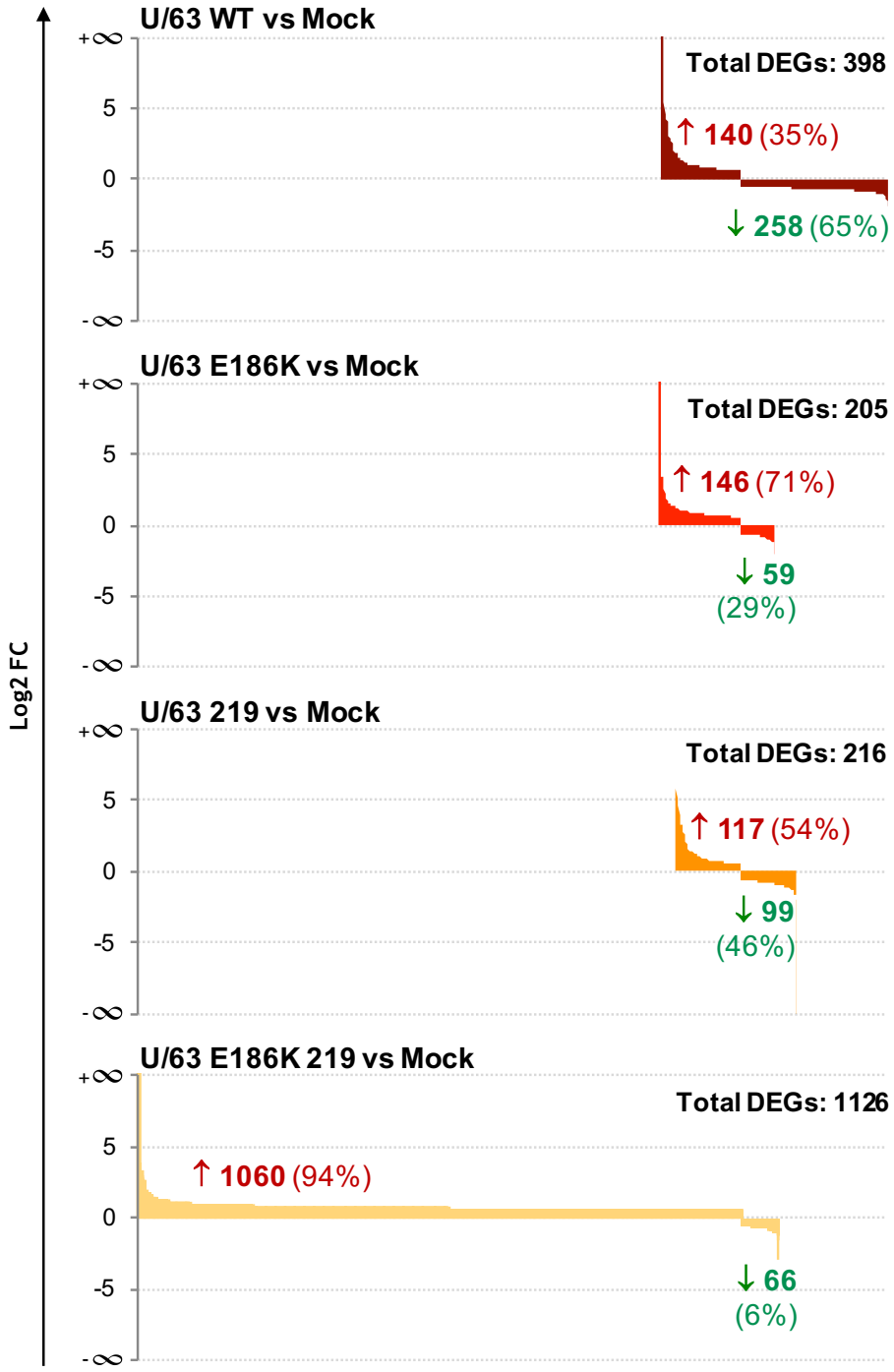


Figure 6-16: Plots representing DEGs compared to mock in U/63-WT- and mutant-infected cells.

Each plot displays the range of host genes found significantly differentially expressed compared to mock-infected cells (Log2FC≥0.58 and p value<0.05) in cells infected with each virus. The total number of DEGs in each condition is indicated in black, up-regulated genes in red and down-regulated genes in green.

The entire list of DEGs for each sample is presented in Appendices 46 to 67.

The analysis of gene sets overlap (Figure 6-17) revealed that U/63-E186K-219 virus modified the expression of the largest number of cellular genes, which were not shared with other conditions (894 up- and 19 down-regulated); whereas the gene expression pattern of E186K single mutant virus-infected cells was the least distinct from other conditions, with 18 up- and 31 down-regulated genes in this condition only. In U/63-WT-infected cells, the majority of genes were alone down-regulated, most of which were not shared with other conditions. Finally, upon U/63-219 infection, a balance between up- and down-regulated genes was observed, with most DEGs not shared with other conditions.

One gene was down-regulated in U/63-E186K-infected cells and up-regulated in U/63-219-infected cells: aldehyde dehydrogenase 1 family member A3 (ALDH1A3), with Log2FC=-1.138555 and 0.652271, respectively. Aldehyde dehydrogenases oxidize aldehydes using either NAD or NADP as co-enzymes, and play important roles in numerous physiological and pathological processes ((Muzio et al., 2012). ALDH1A3 is notably important in the detoxification of exogenous toxicants, risk factors for pulmonary diseases, such as environment pollutant and lipid peroxidation products (Liu et al., 2010).

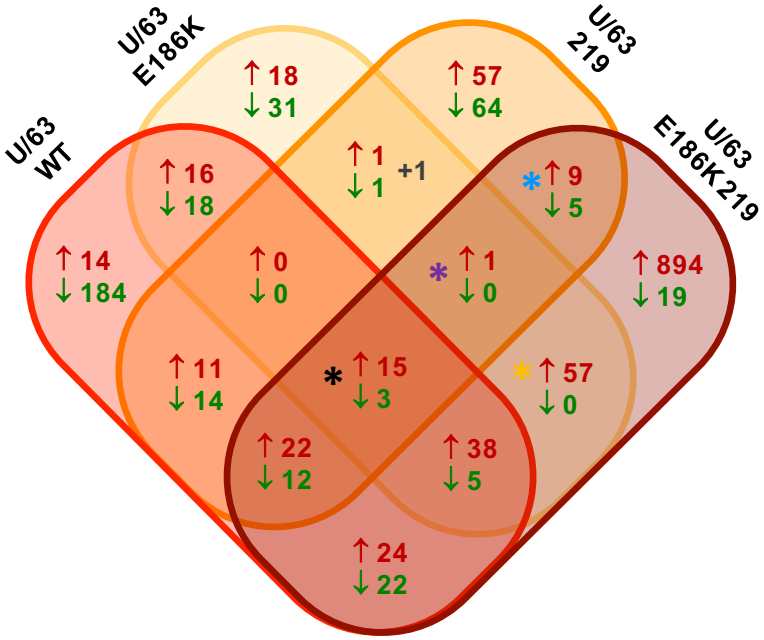


Figure 6-17: Venn diagram representing DEGs compared to mock in common between U/63-infected samples.

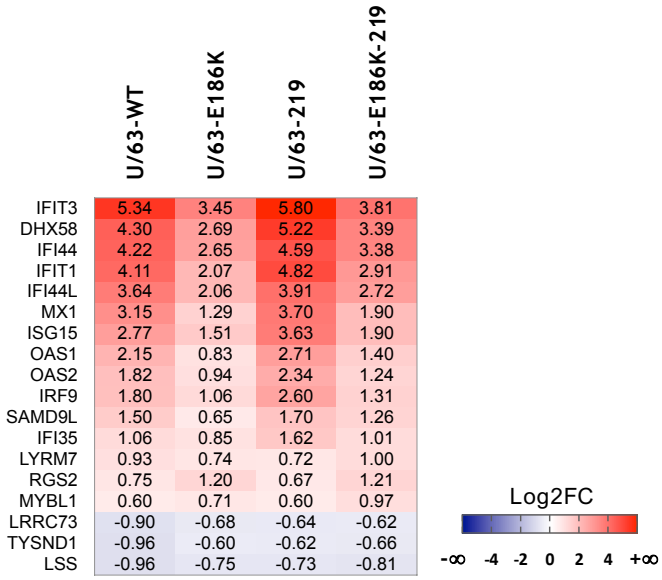
DEGs compared to mock in U/63-infected cells with a Log2FC≥0.58 and p value<0.05 were used this study. Up-regulated and down-regulated genes are indicated in red and green, respectively, while a black number indicates a gene differentially expressed in opposite direction in two different conditions. Asterisks indicate four primary groups of DEGs: black asterisk (*) indicates DEGs of primary group U1, purple asterisk (*) indicates DEGs of primary group U2, yellow asterisk (*) is for DEGs of primary group U3, and blue asterisk (*) are DEGs of primary group U4.

As for O/03 gene sets analysis (Figure 6-7), four primary groups were identified: **(U1)** genes whose expression was similar between WT and mutant virus infections (black asterisk), **(U2)** genes whose expression was similar between mutant viruses, but absent in WT virus infected cells (purple asterisk), **(U3)** genes whose expression was differentially regulated compared to mock in E186K-containing mutant virus-infected cells (yellow asterisks), and **(U4)** genes whose expression was differentially expressed compared to mock in 230-containing mutant virus-infected cells exclusively (blue asterisks).

In group U1 (Figure 6-18), fifteen genes were up-regulated and three were down-regulated. The vast majority of up-regulated genes were involved in the type I IFN pathway: IFIT3, LGP2, IFI44, IFIT1, IFI44L, Mx1, ISG15, OAS1, OAS2, IRF9, SAMD9L, IFI35, viperin. These genes were strongly up-regulated in all infected samples, and more so in U/63-WT- and U/63-219-infected cells.

The MYB proto-oncogene like 1 gene (MYBL1), involved in regulation of RNA pol II promoter-dependent gene transcription, was also up-regulated, and so was LYRM7, previously mentioned in section 6.2.2.4 (group O1, Figure 6-7).

Among down-regulated genes, two were previously found down-regulated in group O1: LSS and TYSND1. Furthermore, leucine rich repeat containing 73 (LRRC73), whose function remains un-explored, was also down-regulated.

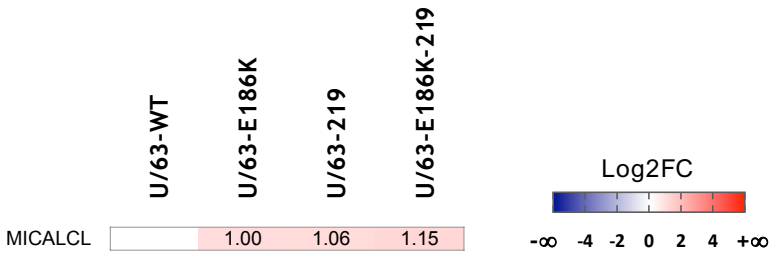


Gene	Key words/functions
IFIT3	Defense response to virus and cellular response to type I IFN, negative regulation of cell proliferation
DHX58	Negative and positive regulation of MDA-5 and RIG-I signaling pathways
IFI44	Defense response to virus
IFIT1	Defense response to virus and cellular response to type I IFN. Intracellular transport of viral protein in host cell, negative regulation of protein binding and helicase activity
IFI44L	Defense response to virus
MX1	Defense response to virus, negative regulation of viral genome replication
OAS1	Defense response to virus, negative regulation of viral genome replication, glucose homeostasis
OAS2	Defense response to virus, RNA catabolic process, protein glycosylation
IRF9	Transcription factor activity, type I IFN signaling pathway
SAMD9L	Endosomal vesicle fusion
IFI35	Type I IFN signaling pathway
LYRM7	Component of mitochondrion, unknown function
RGS2	Negative regulation of G-protein coupled receptor protein signaling pathway, regulation of adrenergic receptor signaling pathway, negative regulation of MAP kinase activity
MYBL1	Regulation of transcription from RNA pol II promoter
LRRC73	Unknown function
TYSND1	Regulation of fatty acid beta-oxidation
LSS	Steroid biosynthetic process, cholesterol biosynthetic process, regulation of protein stability

Figure 6-18: Heatmap highlighting DEGs compared to mock in U/63-WT and mutant virus infected-cells.

Genes whose expression is statistically over-represented (up- or down-regulated with Log2FC≥0.58 and p value<0.05) in U/63-WT- and mutant virus infected-cells compared to mock-infected cells.

In group U2 (Figure 6-19), only one gene was found: MICAL C-terminal like gene (MICALCL), which encodes for a protein that binds mitogen-activated protein kinases (MAPKs) and affects multiple biological processes, such as cell differentiation, and organism development (Miura, 2008).



Gene	Key words/functions
MICALCL	Mitogen-activated protein kinase binding, cell differentiation

Figure 6-19: Heatmap highlighting DEGs compared to mock in U/63 mutant virus infected-cells.

Genes whose expression is statistically over-represented (up- or down-regulated with Log2FC≥0.58 and p value<0.05) in U/63 mutant virus infected-cells compared to mock-infected cells.

In group U3 (Figure 6-20) a large number of up-regulated genes were observed (57 DEGs in total), while no down-regulated genes were found. Among up-regulated genes, several were involved in immune responses, such as phosphodiesterase 4B (PDE4B) an enzyme induced by LPS (Ma et al., 1999) (Selige et al., 2011, Jin et al., 2010); Interleukin 13 receptor alpha 2 gene (IL13RA2) that encodes for the receptor of IL-13 (Daines et al., 2006); Cluster differentiation 80 gene (CD80, also called B7-1), encoding for a molecule classically found on antigen presenting cells where it plays a pivotal in the activation and regulation of T cell function (Chen and Flies, 2013), but interestingly has also been found on fibroblasts (Pechhold et al., 1997, Corrigall et al., 2000); Matrix metallopeptidase 1 gene (MMP1), encoding one of the 23 family members of calcium and zinc dependent enzymes implicated in physiological and inflammatory pathological processes of the airways and lung parenchyma, such as asthma, chronic obstructive pulmonary disease (COPD) and emphysema (Cataldo et al., 2003); Sphingosine-1-phosphate receptor 1 gene (S1PR1) encoding a G protein-binding receptor , which plays important immune functions, such as promoting pro-inflammatory responses, T cell trafficking and development (Chi, 2011), and diverse physiological processes (Rivera et al., 2008, Aoki et al., 2016); Oncostatin M receptor gene (OSMR), encoding for the receptor of oncostatin-M (OSM), an IL-6 family member able to stimulate the eosinophil-selective CC chemokine eotaxin-1 *in vitro*, and eosinophil

accumulation in mouse lung in vivo (Hermanns, 2015); Kit ligand gene (KITLG) that encodes a pleiotropic factor thought to impact on cell development, hematopoiesis, and mast cell proliferation disorders (Cruse et al., 2014); COMM domain containing 8 gene (COMMD8) encodes a protein that interacts with the coiled-coil domain-containing protein 22 (CCDC22) and regulate the turnover of I-kappa-B and activation of NF-kappa-B, a master regulator of inflammation (Starokadomskyy et al., 2013); the NF-κB repressing factor gene (NKRF), encoding a DNA-binding nuclear transcriptional repressor of NF-κB, which has been associated with suppression of IP-10 and IL-8 in alveolar macrophages during pulmonary tuberculosis (Huang et al., 2013); the β-1,3-N-acetylglucosaminyltransferase 2 gene (B3GNT2), which encodes for a poly-lactosamine synthase that generates putative immune regulatory factors, presumably suppressing excessive responses during immune reactions (Togayachi et al., 2007); Leucine rich repeat protein 1 gene (LRR1), which encodes a protein that specifically interacts with and negatively regulates member 9 of the tumor necrosis factor receptor (TNFR) superfamily (TNFRSF9, also called 4-1BB), which result in the activation of NF-κB and JNK1 (Jang et al., 2001); and Glutaredoxin 2 gene (GLRX2) encodes an antioxidant enzyme that facilitates the maintenance of mitochondrial redox homeostasis (Reynaert et al., 2007), and more importantly the Glutaredoxin/S-glutathionylation axis regulates interleukin-17A-induced pro-inflammatory responses in lung epithelial cells (Nolin et al., 2014)

Some genes were involved in fibroblasts differentiation and migration and tissue remodelling, such as Hyaluronan synthase 2 gene (HAS2), which encodes the enzyme producing hyaluronan, a major component of extracellular matrix secreted by airway structural cells (Walker et al., 2017) and important in fibroblast to myofibroblast differentiation (Midgley and Bowen, 2015); Integrin subunit alpha 10 gene (ITGA10), encoding a protein of the integrin alpha family, which are transmembrane adhesion molecules mediating cell-cell and cell-extracellular matrix attachment, and also regulate cell growth, proliferation, migration and apoptosis (Lenci et al., 2012); Plasminogen receptor with a C-terminal lysine gene (PLGRKT) that encodes for the plasminogen receptor and favours its activation to plasmin (Herren et al., 2003), which in turn is important for extracellular matrix remodelling and cell migration; ADAM metallopeptidase with thrombospondin type 1 motif 1 gene (ADAMTS1), encoding a protein that belongs to the ADAMTS protein family and plays a role in extracellular matrix proteoglycan (versican) processing (Cross et al., 2005, Cal et al., 2002); Neuroepithelial cell transforming 1 gene (NET1), which encodes the NET1 protein, a Ras homolog family member A-specific guanine nucleotide exchange factor induced by TGF-β and controls the formation of stress fibers and reorganisation of cytoskeletal structures (Shen et al., 2001); Tensin 4 gene (TNS4, also

called cten) encoding one of the four members of the Tensins family (Lo, 2004), cytoplasmic proteins that are localized to integrin-linked focal adhesions and weaken cellular adhesion and increases cell migratory activity (Mouneimne and Brugge, 2007); Ubiquitin associated and SH3 domain containing B gene (UBASH3B, also called p70S6K) (Pei et al., 2006), which encodes a protein that inhibits endocytosis of epidermal growth factor receptor (EGFR) and platelet-derived growth factor receptor (Lee et al., 2013); Cellular repressor of E1A stimulated genes 2 (CREG2) that encodes for a secreted glycoprotein inhibiting cell proliferation and enhancing differentiation of pluripotent stem cells (Kunita et al., 2002); Follistatin gene (FST), which encodes for an angiogenic protein (Frezzetti et al., 2016) that binds and inhibits activins, members of the transforming growth factor beta superfamily that play a significant role in lung regeneration (George et al., 2017, Zhao et al., 1996); and the sprouty RTK signaling antagonist 2 gene (SPRY2) that encodes a protein known to downregulate angiogenesis during wound healing (Wietecha et al., 2011).

Other up-regulated genes from this group were involved in DNA stability and gene expression, such as Helicase lymphoid specific gene (HELLS) encoding a novel member of SNF2 family (Yano et al., 2004), which possess an helicase and ATPase activity and is thought to regulate transcription of certain genes by altering the chromatin structure around those genes (von Eyss et al., 2012). The endothelial cell specific molecule 1 gene (ESM1, also called Endocan) that encodes an endothelial cell-associated proteoglycan preferentially expressed by renal and pulmonary endothelium, which is upregulated by pro-angiogenic and pro-inflammatory cytokines and reflects endothelial activation and dysfunction (Kechagia et al., 2016); GEN1 gene, a member of the Rad2/XPG nuclease family shown to promote Holliday junction (HJ) resolution (Chan and West, 2015), which link sister chromatids or homologous chromosomes and are essential for the repair of stalled/collapsed replication forks in mitotic cells (Rass et al., 2010); the DNA replication helicase/nuclease 2 (DNA2) gene, which encodes a conserved helicase/nuclease involved in the maintenance of mitochondrial and nuclear DNA stability (Budd and Campbell, 2013); the MMS22L gene that encodes a protein interacting with the scaffold-like protein Nfkbil2 (Piwko et al., 2010), and forms a complex that contributes to genome stability by regulating the chromatin state at stalled replication forks (O'Connell et al., 2010); SMC5-SMC6 complex localization factor 1 gene (SLF1) whose encoded-protein forms a complex with SLF2 and RAD18 involved in DNA post-replication repair and homologous recombination repair (Liu et al., 2012, Raschle et al., 2015); Origin recognition complex gene (ORC) that encodes a protein involved in the initiation of chromosomal replication (Gibson et al., 2006); Establishment of sister chromatid cohesion N-acetyltransferase 2 gene (ESCO2) encoding a protein belonging to the highly conserved Eco1/Ctf7 family involved in establishment of sister chromatid cohesion

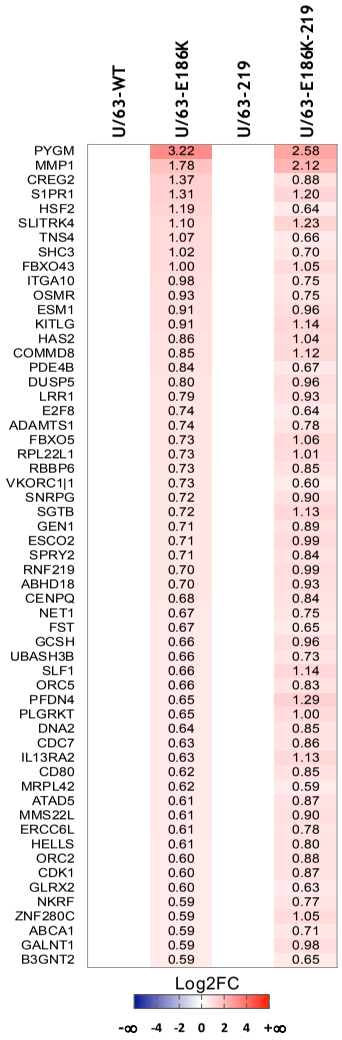
during S phase and has acetyltransferase activity (Nishihara et al., 2010); Centromere protein Q gene (CENPQ) encoding a subunit of the centromeric complex required for proper kinetochore function and mitotic progression (Okada et al., 2006); ATPase family AAA domain containing 5 gene (ATAD5, also known as ELG1), which encodes a protein playing an important role in S phase to preserve genomic stability (Maleva Kostovska et al., 2016, Shiomi and Nishitani, 2013), and is induced during recovery from a variety of DNA damage (Sikdar et al., 2009); and Small nuclear ribonucleoprotein polypeptide G gene (SNRPG) encoding a component of the U1, U2, U4, and U5 small nuclear ribonucleoprotein complex, precursors of the spliceosome (Urlaub et al., 2001).

Several genes were involved in regulation of the cell cycle and cell survival, such as the cell division cycle 7 (CDC7) gene, which encodes a cell division cycle protein with kinase activity critical for the G1/S transition (Datta et al., 2017); Excision repair cross-complementation group 6 like (ERCC6L), a gene encoding for a DNA helicase of the SNF2 family (Xu et al., 2005, Pu et al., 2017); Heat shock transcription factor 2 gene (HSF2) encoding a protein that belongs to the HSF family of transcription factors and binds specifically to heat-shock promoter element and activate transcription and cellular differentiation in response to cellular stresses (Tateishi et al., 2009); E2F transcription factor 8 gene (E2F8), encoding a novel identified E2F family member, which controls the mammalian cell cycle (Ye et al., 2016); Cyclin dependent kinase 1 gene (CDK1, formerly called cdc2) encoding an ISG (Sze et al., 2013) that controls entry into S- and M-phase (Heim et al., 2017); RB binding protein 6, ubiquitin ligase (RBBP6) interacts with both p53 and Rb (Dlamini et al., 2016), and plays a role in apoptosis induction and regulation of the cell cycle (Pugh et al., 2006); Ribosomal protein L22 like 1 gene (RPL22L1) encoding a minor ribosomal component belonging to the ribosomal protein L22e family, ubiquitously expressed and whose deletion induces p53 expression and cell death of α/β-lineage T lymphocytes specifically; F-box protein 5 (FBXO5) and 43 (FBXO43) genes encoding E3 ubiquitin ligase receptors (Jin et al., 2004), shown to regulate the G(1)/S transition of the eukaryotic cell cycle (Chiaur et al., 2000); Dual specificity phosphatase 5 gene (DUSP5), which encodes an inducible nuclear protein (Temple et al., 2001) ubiquitously expressed that specifically interacts with and deactivates extracellular signal-regulated kinases (ERK1 and ERK2) responsible for cell proliferation, differentiation, and survival (Talipov et al., 2016); and SHC adaptor protein 3 encoding gene (SHC3) which encodes a member of the Shc family of multifunctional docking proteins involved in Ras-extracellular signal-regulated kinase (ERK) and the PI3K/Akt pathways (Conti et al., 2001).

Two genes were involved in synthesis and bioavailability of lipids: ABCA1 gene, which has previously been found in group O4 (Figure 6-11); and α/β hydrolase domain containing 18 gene (ABHD18), whose function is unknown (Thomas et al., 2014).

Two genes were involved in protein folding: Small glutamine rich tetratricopeptide repeat containing beta gene (SGTB), induced by IL-1β (Bao et al., 2016); and Prefoldin subunit 4 gene (PFDN4) (Hartl and Hayer-Hartl, 2002).

Finally, several genes were involved in metabolic processes and minerals homeostasis, such as the zinc finger protein 280C gene (ZNF280C) encoding a member of the Cys2His2 zinc finger domain-containing protein family (Gamsjaeger et al., 2007) extremely abundant in higher eukaryotes and which seem to act as redox-sensitive molecular switches controlling crucial cellular processes (Kroncke and Klotz, 2009); Ring finger protein 19 gene (RNF219), ubiquitously expressed in various tissues (Huh et al., 2008), and encodes a RING finger type E3 ubiquitin ligase, which binds the intracellular carboxyl terminus of Calcium-sensing receptors, and regulate calcium homeostasis (Huang et al., 2006); Polypeptide N-acetylgalactosaminyltransferase 1 gene (GALNT1), which encodes a member of the family of glycosyltransferases (GalNAc-Ts) responsible for the initiation of mucin-type O-glycosylation (Tian et al., 2015); Glycine cleavage system protein H gene (GCSH) encodes one of the four mitochondrial protein components of the glycine cleavage system (Kure et al., 2001); Mitochondrial ribosomal protein L42 gene (MRPL42), which encodes a mitochondria-specific protein helping in protein synthesis within the mitochondrion (Lightowlers et al., 2014); Muscle glycogen Phosphorylase gene (PYGM) encoding a protein that catalyses the degradation of glycogen to glucose-1-phosphate (Nam et al., 2011); SLIT and NTRK like family member 4 gene (SLITRK4), which belongs to the SLITRK family, along with the axonal growth-controlling protein SLIT and neurotrophin receptors (Aruga et al., 2003); Vitamin K epoxide reductase complex subunit 1 gene (VKORC1), which encodes the catalytic subunit of the vitamin K epoxidase reductase, complex responsible for the reduction of inactive vitamin K 2,3-epoxide to active vitamin K (Bodin et al., 2005).



Gene	Key words/functions
PYGM	Carbohydrate metabolic process, glycogen catabolic process
MMP1	Proteolysis, positive regulation of protein oligomerization
CREG2	Oxidation-reduction process
S1PR1	Positive regulation of cell proliferation, migration, regulation of cell adhesion, leukocyte chemotaxis, actin cytoskeleton reorganization, T cell migration
HSF2	Positive regulation of transcription from RNA pol II promoter
SLITRK4	Axonogenesis, positive regulation of synapse assembly
TNS4	Protein localization
SHC3	Intracellular signal transduction
FBXO43	Metal ion binding
ITGA10	Cell-matrix adhesion, integrin-mediated signaling pathway
OSMR	Positive regulation of cell proliferation, response to cytokine
ESM1	Positive regulation of cell proliferation
KITLG	Positive regulation of leukocyte migration, of mast cell proliferation, of hematopoietic stem cell proliferation, of myeloid leukocyte differentiation, negative regulation of mast cell apoptotic process
HAS2	Vasculogenesis, cell adhesion, positive regulation of cell proliferation and monocyte aggregation, hyaluronan metabolic process, positive regulation of cell migration, cellular response to IL-1 and TNF
COMMD8	Regulation of DNA-templated transcription
PDE4B	Cellular response to LPS, neutrophil chemotaxis, positive regulation of IL-2 and IFN-γ production
DUSP5	Inactivation of MAPK activity, protein dephosphorylation
LRR1	Signal transduction
E2F8	Negative regulation of transcription from RNA pol II promoter, negative regulation of cytokinesis
ADAMTS1	Endopeptidase activity, proteolysis
FBXO5	Regulation of mitotic cell cycle, vesicle organization, microtubule polymerization
RPL22L1	Structural constituent of ribosome and protein translation
RBBP6	Ubiquitin-protein transferase activity, zinc ion binding
VKORC1	Catalytic subunit of the vitamin K epoxide reductase complex
SNRPG	mRNA splicing
SGTB	Protein homo- & hetero-dimerization
GEN1	DNA double-strand break repair
ESCO2	Regulation of DNA replication, DNA double-strand break repair
SPRY2	Negative regulation of cell proliferation and apoptotic process, positive regulation of gene expression, lung development

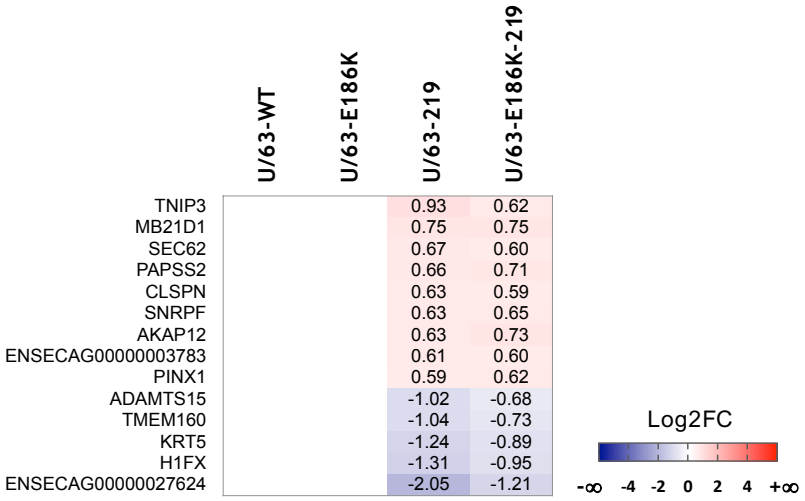
Gene	Key words/functions
RNF219	Metal ion binding, unknown function
ABHD18	Unknown function
CENPQ	Positive regulation of protein localization to kinetochore
NET1	Positive regulation of apoptotic process
FST	Negative regulation of transcription from RNA pol II promoter, hematopoietic progenitor cell differentiation
GCSH	Protein lipoylation
UBASH3B	Negative regulation of protein kinase activity, regulation of release of sequestered calcium ion into cytosol
SLF1	Cellular response to DNA damage stimulus
ORC5	DNA replication initiation
PFDN4	Protein folding, chaperone binding
PLGRKT	Chemotaxis, inflammatory response
DNA2	DNA replication checkpoint, telomere maintenance
CDC7	Positive regulation of G2/M transition of mitotic cell cycle, DNA replication association, positive regulation of cell proliferation
IL13RA2	Negative regulation of immunoglobulin production, immunoglobulin-mediated immune response, negative regulation of mast cell degranulation
CD80	Cellular response to lipopolysaccharide, T cell co-stimulation, positive regulation of T cell proliferation
MRPL42	Mitochondrial
ATAD5	Mitochondrial translational processes
MMS22L	Double-strand break repair
ERCC6L	Cell cycle and division
HELLS	DNA methylation, negative regulation of apoptotic process, lymphocyte proliferation
ORC2	DNA replication
CDK1	Microtubule cytoskeleton organization, DNA repair, positive regulation of G2/M transition of mitotic cell cycle, histone phosphorylation, negative regulation of apoptotic process
GLRX2	Response to hydrogen peroxide, cell redox homeostasis
NKRF	Positive regulation of transcription from RNA pol II promoter
ZNF280C	Regulation of DNA-templated transcription
ABCA1	Cholesterol metabolic process, lipid transport,
GALNT1	Protein glycosylation, glycosyl group-transferase activity
B3GNT2	Glycosyl group-transferase activity

Figure 6-20: Heatmap highlighting DEGs compared to mock in U/63-E186K-containing virus infected-cells.

Genes whose expression is statistically over-represented (up- or down-regulated with Log2FC≥0.58 and p value<0.05) in U/63-K186E-containing virus infected-cells compared compared to mock-infected cells.

Finally, group U4 (Figure 6-21) counted 9 up-regulated and 5 down-regulated genes. Among up-regulated genes was TNFAIP3 interacting protein 3 (TNIP3, also known as ABIN-3), an IL-10-induced gene product expressed by negative regulation of TNF receptor- and TLR4-induced NF-κB activation (Verstrepen et al., 2008); Mab-21 domain containing 1 gene (MB21D1, also known as cyclic GMP-AMP synthase (cGAS)) that encodes a cytosolic sensor of dsDNA, which induces an antiviral state and the secretion of type I IFN (Ablasser et al., 2013, Mackenzie et al., 2017); Sec62 gene that encodes a critical constituent of the translocon complex, which is important for the maintenance of endoplasmic reticulum homeostasis (Fumagalli et al., 2016); the 3'-phosphoadenosine 5'-phosphosulfate synthase 2 gene (PAPSS2), which encodes an enzyme that synthesizes high energy sulfate donor (PAPS) used in all sulfonation reactions (Schroder et al., 2012); Claspin protein-encoding gene (CLSPN), which controls cell cycle arrest in response to DNA damage (Chini and Chen, 2003, Chini and Chen, 2004); Small nuclear ribonucleoprotein polypeptide F gene (SNRPF) encoding one of the seven Sm proteins constituent of the nuclear pre-mRNA processing machinery and involved in pre-mRNA splicing (Kambach et al., 1999); A-kinase anchoring protein 12 gene (AKAP12) encoding a scaffold protein that associates with protein kinases A and C, and plays a role in signal transduction (Chen and Malbon, 2009), regulation of cell migration (Lee et al., 2011), cell cycle progression and suppression of apoptosis (Yoon et al., 2007); and PIN2/TERF1 interacting telomerase inhibitor 1 gene (PINX1) encoding a telomerase inhibitor (Cheung et al., 2012) implicated in telomere maintenance (Yoo et al., 2014). In addition, one novel pseudogene whose function is unknown was also up-regulated in this group (ENSECAG00000003783).

Among down-regulated genes was the ADAM metallopeptidase with thrombospondin type 1 motif 15 gene (ADAMTS15), encoding a protein that belongs to the ADAMTS protein family and plays a role in extracellular matrix proteoglycan processing (Cross et al., 2005, Cal et al., 2002); Transmembrane protein 160 gene (TMEM160) encoding a component of the mitochondrion, whose function is unknown; Keratin 5 gene (KRT5) expressed by lung progenitor cells (krt5+) important in lung remodelling (Xi et al., 2017) and regeneration of epithelial cells following an influenza virus infection (Yee et al., 2017, Quantius et al., 2016). H1 histone family member X gene (H1FX, also called NCAMP-1) encoding the least characterized member of the H1 family (Ichihara-Tanaka et al., 2017), which is ubiquitously expressed (Evans et al., 2009) and has important functions in mitotic progression (Takata et al., 2007). Finally, ENSECAG00000027624 is a non-coding RNA, whose function is unknown.



Gene	Key words/functions
TNIP3	Positive regulation of transcription from RNA pol II promoter, TLR4 signaling pathway
MB21D1	Defense response to virus, positive regulation of type I IFN production
SEC62	Post-translational protein targeting to endoplasmic reticulum membrane
PAPSS2	Sulfate assimilation
CLSPN	G2 DNA damage checkpoint
SNRPF	mRNA splicing via spliceosome
AKAP12	Positive regulation of protein kinase A and C signaling, signal transduction
ENSECAG00000003783	Novel pseudogene, unknown function
PINX1	Maintain telomere integrity by regulating TRF1
ADAMTS15	Metallopeptidase activity
TMEM160	Integral component of mitochondrion, unknown function
KRT5	Structural molecule activity, scaffold protein binding
H1FX	Nucleosome assembly
ENSECAG00000027624	Non-coding RNA, unknown function

Figure 6-21: Heatmap highlighting DEGs compared to mock in U/63-219-containing virus infected-cells.

Genes whose expression is statistically over-represented (up- or down-regulated with Log2FC≥0.58 and p value<0.05) in U/63-219-containing mutant virus infected-cells compared to mock-infected cells.

6.2.3.5. Top canonical signalling pathways modified by U/63 viruses

Finally, as performed for O/03-infected cells (Figure 6-11), the datasets of DEGs identified in U/63-WT and mutant virus-infected cells were imported into the IPA tool. Figure 6-22 represents a hierarchical clustering heatmap showing the list of activated canonical pathways after U/63-WT and mutant virus infection.

Very few pathways were significantly modified in U/63 single mutant virus-infected cells, and the largest number of activated pathways were observed for U/63-E186K-219-infected conditions, closely followed by U/63-WT-infected cells. In addition, the majority of pathways were negatively-regulated in U/63-WT-infected samples, while being positively-regulated in cells infected with U/63 double mutant virus, e.g. inflammatory pathways with positive z-score for U/63-double mutant virus and negative z-score for WT (FcεRI signalling, IL-17A signalling in airway cells, IL-6 signalling, LPS-stimulated MAPK signalling, CXCR4 signalling).

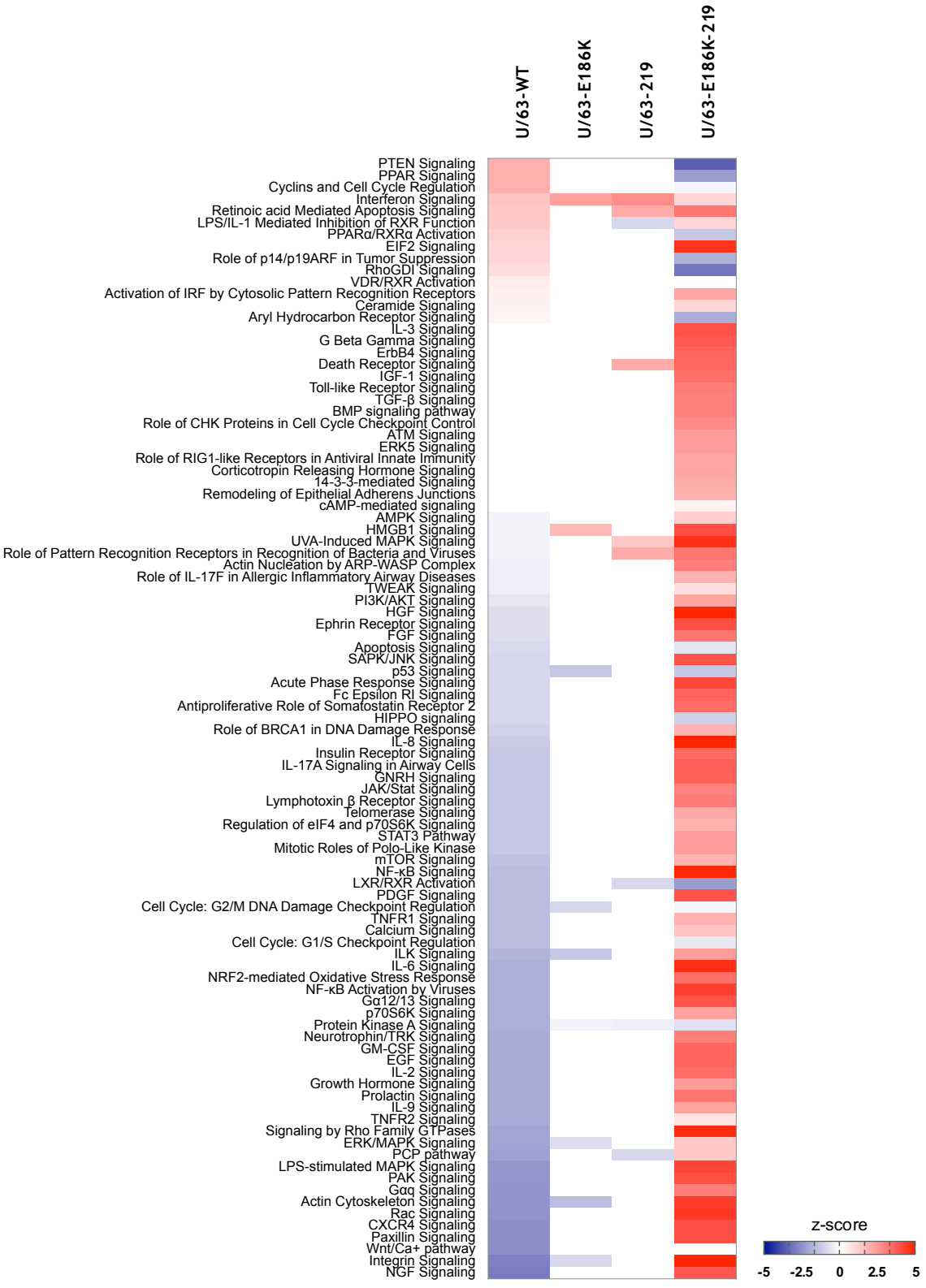


Figure 6-22: Top canonical signalling pathways activated by U/63-WT and mutant viruses.

As for Figure 6-12, the Ingenuity Pathway Analysis (IPA) tool was used to generate a list of significant canonical pathways activated in virus-infected cells compared to mock-infected cells.

6.3. Discussion

In this chapter, high-throughput RNA sequencing technologies were used to analyse the host response to two phylogenetically distinct EIVs, O/03 and U/63. In addition, the impact of two NS1 evolutionary markers on virus-host interaction, E186K substitution and C-terminal truncation, was evaluated by comparing the transcriptional profile of equine cells infected with WT and NS1 186 and C-terminus mutant viruses.

To prepare the samples for sequencing, E.Derm cells were infected with each virus or mock infected for 8h, which corresponds to the end of the eclipse phase in these cells (seen in the previous chapter, Figure 5-7). The MOI was also adjusted for each independent infection to ensure that a similar number of cells were infected in all conditions, which was checked by flow cytometry (Figure 6-1). In doing so, variations in levels of infection between samples were limited. However, both for O/03- and U/63-infected samples the majority of cells (approximately 70%) whose transcriptome was analysed was not infected. Thus, the transcriptomic results should be interpreted as a result of the cellular response to infection in foci of infection as well as the cellular response of cells surrounding these foci. Although this situation is likely to arise during a natural infection (Pociask et al., 2017), the relevance of RNAseq experiment in cells with such a low level of infection is debatable. In addition, using a virus stock generated from a cell line without purifying to suspension to discard any soluble factors secreted by infected cells (e.g. cytokines, growth factors, etc.) would have been preferable. To obtain a higher percentage of infection and a purer viral suspension the use of a density gradient (e.g. conventional gradient of sucrose and a gradient of iodixanol) (Hutchinson & Fodor, 2014) coupled with centrifugal inoculation (or spinoculation) (Guo et al., 2011, Chaipan et al., 2009) would have been preferable.

In this work, three biological replicates were used for each condition, and differentially expressed genes compared to mock were considered statistically significant with fold change superior or equal to 1.5 (Log2FC≥0.58). In addition, the total number of reads for each sample was relatively high with a mean of 41.2 million reads for mock- and O/03-infected cells (Figure 6-4), and a mean of 35.2 million reads for U/63-infected cells (Figure 6-14). The RNA quality for all samples was also high, with RINs above 9.5 on a scale from 1 to 10 (Figure 6-2). According to a survey of best practices for RNA sequencing data analysis (Conesa et al., 2016), these parameters gave a statistical power of 43% (Consortium, 2013). The statistical power to detect differential gene expression varies with effect size (fold change), sequencing depth and number of replicates. To reach 100% of statistical

power, ten biological replicates per condition and a fold change of 2 would have been necessary.

Overall, all biological replicates were closely distributed (Figure 6-3 & 6-13), which indicated a low variability between each independent experiment. In addition, O/03-WT and K186E-containing mutant virus-infected cells were closely clustered, while O/03-230-infected cells were clustering with mock-infected cells. Similarly, U/63-WT and U/63-E186K-containing mutant virus infected cells were closely clustered, while U/63-219-infected cells were more closely distributed to mock-infected cells. The underlying reasons remain unclear, however, it is possible that modifying the NS1 tail alone alters virus-host interaction more strongly than modifying NS1 codon 186.

Unsurprisingly, the large majority of reads aligned to the host genome (Equus caballus 2, GCA_000002305.1) in all samples, with 36.6 million reads in O/03-infected cells and 32.3 million reads in U/63-infected cells (Figure 6-4 and 6-14, respectively). Only a small fraction of reads aligned to the viral genome, with a mean of 3.5 million of reads in O/03-infected cells (Figure 6-4) and a mean of 0.4 million of viral genome reads in U/63-infected cells (Figure 6-14). The level of infection was slightly higher for O/03 viruses (23.1% to 31.6%) (Figure 6-1A) than for U/63 viruses (16.8% to 29.0%) (Figure 6-1B), thus fewer viral reads were expected in U/63-infected conditions than in O/03-infected conditions. However, this difference could also be explained by variations in polymerase efficiency between O/03 and U/63 virus. This idea is further supported by results obtained with mini-replicon technologies showing differences in polymerase efficiency between U/63 and O/03 in human cells (Coburn, 2017). These results also correlate with the lack of replication of U/63 in equine cells compared to O/03 (Figure 5-10 & 5-11).

The characterization of virus-host interaction confirmed the importance of NS1 amino acid 186 and the C-terminal tail on viral control of cellular and viral gene expression. Indeed, mutations introduced in U/63 and O/03 NS1 altered viral gene expression. Indeed, U/63-WT expressed significantly more PB2 and NA genes than the corresponding isogenic mutant viruses, and the truncation of U/63 NS1 led to an increased in expression of the M gene (Figure 6-15). In the case of U/63, no synergistic effect was found between mutation of codon 186 and 220. In contrast, for O/03 the concomitant introduction of K186E substitution and NS1 C-terminus extension increased viral expression of HA as well as NA and NS, and decreased expression of NP (Figure 6-5). Interestingly, Park et al. showed that an increased expression of HA, NA and NS were associated with the ability to replicate with

better efficiency at earlier times post infection (Park et al., 2015), which correlated with the data collected in Chapter 5 (Figure 5-7 & 5-11). Thus, it seemed that for a contemporary virus like O/03, the reversion of NS1 adaptive markers strongly disrupted viral gene expression, and their introduction in an emergent virus like U/63 had minimal effects.

The total number of differentially expressed genes compared to mock in WT virus-infected cells were comparable between O/03 and U/63 (429 and 398 DEGs, respectively) (Figure 6-6 & 6-16). However, close to 60% of DEGs were up-regulated in O/03-WT-infected cells, while 65% of DEGs were down-regulated in U/63-WT-infected cells. This indicated that the strategy of control of infected cells were fundamentally different between O/03 and U/63. Indeed, the equine-adapted virus O/03 promoted up-regulation of host genes, while the emergent U/63 virus favoured a downregulation. Finally, both U/63 single mutants seemed strongly attenuated in equine cells, as very few cellular pathways were modified during infection. In addition, the changes introduced in NS1 seemed to have impacted on U/63's ability to shutdown gene expression, as the total number of DEGs were strongly reduced and the majority of genes were up-regulated (Figure 6-16). This seems to indicate that U/63 requires E186 and a full length NS1 to shutoff host gene expression, and the introduction of an evolutionary markers in NS1 in the context of an emergent virus only reduced viral fitness.

Moreover, the gene expression pattern of cells infected with O/03-WT was highly distinct from that of cells infected with O/03 mutant viruses (Figure 6-7). In contrast, in the context of U/63 infection, the virus inducing the most distinct cellular gene expression pattern was the double mutant virus (Figure 6-17). In addition, the introduction of E186 in O/03 NS1 favoured the downregulation of genes, while the extension of O/03 NS1 tail increased the percentage of up-regulated genes. Consistent with these results, introducing both changes in O/03 NS1 mitigated this effect (Figure 6-6). In the context of U/63 however, introduction of K186 in NS1 or the truncation of the protein increased the number of up-regulated genes. Taken together these data further supports the idea that E186 plays an important role in the virus-induced host gene expression shutoff during infection, while a short NS1 protein harbouring K186 favours cellular gene expression and the activation of a large number of cellular pathways.

Infection of equine cells with the WT version of the mammalian-adapted O/03 virus was also associated with up-regulation of a very small number of ISGs (IFIT3 and IFIT1) (Figure 6-8). In contrast, all O/03 isogenic mutants up-regulated a large range of ISGs (e.g.

LGP2, IFI44, ISG15, USP41, Mx1, OASL2, OAS2, IRF9, IFI35) (Figure 6-9). In addition, the virus the least capable of controlling ISG induction was the O/03 E186-containing mutant viruses. Indeed, apart from the genes described above, these viruses also up-regulated other PRRs and ISGs (such as RIG-I, MDA5, IFITM3L, PKR, viperin, GBP3, PARP9, SAMD9, and SAMDL9) (Figure 6-10). In contrast, infection of equine cells with WT and mutant U/63 viruses was associated with an up-regulation of a comparable number of ISGs (e.g. IFIT3, LGP2, IFI44, IFIT1, IFI44L, Mx1, ISG15, OAS1, OAS2, IRF9, SAMD9L, IFI35, viperin) (Figure 6-18). Thus, it seemed that a short NS1 harbouring K186 impacts on the ability of a mammalian adapted virus, like O/03, to control the antiviral response of the host at a transcriptional level, however, these markers are not sufficient to grant this ability to an emergent virus, like U/63. These results were in accordance with another study, which analysed the cellular response to infection with an avian influenza virus. These authors showed that infection of mammalian cells with an avian virus increased expression of IFN, ISGs and viral sensors (Park et al., 2015), and this was associated with a limited replication efficiency, as found in with O/03 mutants or U/63 viruses (Chapter 5).

Interestingly, the extension of O/03 NS1 tail led to an up-regulation of the alpha subunit of the receptor for IL-22, a cytokine important for pulmonary repair post IAV infection (Noyama et al., 2017, McAleer and Kolls, 2014). This C-terminal extension of O/03 was also associated with a downregulation of ELOVL6, important for lung homeostasis (Sunaga et al., 2013). This raises the questions of the impact of Influenza virus exposure on lung inflammation and repair, and the implication of NS1 in these mechanisms.

Surprisingly, only one gene (MICALC) was shared between cells infected with U/63 mutant viruses (Figure 6-19), while a large number of up-regulated genes (57 in total) were found in common between cells infected with U/63 E186K-containing viruses. Importantly, these genes were involved in immune responses, fibroblasts differentiation, migration and tissue remodelling, DNA stability, gene expression, regulation of the cell cycle and cell survival, synthesis and bioavailability of lipids, metabolic processes and minerals homeostasis. This tends to suggest that while NS1 C-terminus has only a negligible impact on U/63's control of infected cells, E186 plays a major role in virus-host interaction and is necessary for general gene shutoff upon infection (Figure 6-20).

Additionally, as seen in Figure 6-12, the O/03-WT virus activated a large number of cellular pathways, and notably up-regulated pathways involved in cell survival, such as the PI3K/AKT pathway, and down-regulated pathways associated with cell cycle arrest and cell

death (i.e. apoptosis signalling, telomerase signalling, cell cycle G1/S checkpoint regulation, p53 signalling, role of BRCA1 in DNA damage response, cell cycle G2/M checkpoint regulation). Interestingly, activation of AKT during an influenza virus infection is important to favour viral protein expression and regulate cell death in infected cells (Kuss-Duerkop et al., 2017). Furthermore, activation of death receptor signalling pathways was previously associated with high pathogenicity in vivo (Park et al., 2015). Taken together, this data seems to indicate that O/03 possesses mechanisms to maintain cells alive and favour viral protein production, and a short version of NS1 with K186 may be important for this function. As the phosphorylation status of AKT during an influenza virus infection yields different outcomes (Kuss-Duerkop et al., 2017), it would be interesting to evaluate the different phosphorylation states of AKT in cells infected with O/03-WT and to compare it with the situation upon U/63 infection. Moreover, evaluating the impact of residue 186 and C-terminal tail on AKT phosphorylation status would be interesting.

In summary, combining cellular gene profiling with reverse genetic techniques has given new insights into the cellular responses of equine cells to EIV infection. This approach has also highlighted the strain-specific roles of NS1. Although further work would be necessary to go beyond the current results, the data collected tends to indicate that NS1 evolutionary markers E186K and the C-terminal truncation have a high impact on virus-host interaction and were likely detrimental for EIV adaptation to horses.

Chapter 7

Final reflection and potential for future research

7. Final reflection and potential for future research

The primary aim of this project was to identify molecular markers of adaptation of EIV to horses in the NS1 protein, the main viral antagonist of the host immune response. To aid with this investigation, a phylogenetic analysis was associated with cloning and site-directed mutagenesis techniques, as well as with reporter assays in order to identify markers of evolution of NS1 function. Then, the impact of NS1 evolutionary markers on EIV infection phenotype and virus-host interaction was evaluated using by reverse genetic systems and high-throughput sequencing technologies.

First by analysing the phylogenetic relationship of one-hundred-and-seventy-five EIV NS1 sequences (Chapter 4), and by comparing their amino acid sequences, fifteen residue changes fixed throughout evolution were identified. These changes correlated with previous reports from the literature (Murcia et al., 2011, Murcia et al., 2010, Barba and Daly, 2016). These changes comprised A112T and E186K substitutions in the early 1970s, R44K reversion and R59H, E71K, A86T and V230I substitutions in the early 1990s, and P216S substitution and C-terminal truncation (R220*) in the late 1990s. Although it was beyond the scope of this work, analysing in more detail the impact of each residue on NS1 function, as well as evaluating potential synergistic effects of these amino acids on NS1 function would be interesting.

To move forward with the characterization of EIV NS1, thirteen phylogenetically distinct NS1 proteins spanning the entire EIV lineage were cloned into an expressing vector, and tested in reporter assays (Hale et al., 2010, Kochs et al., 2007, Turkington et al., 2015) to compare their function. It was observed that during the first decade post viral emergence, NS1 strongly repressed general gene expression through a CPSF30-dependent mechanism, which required the presence of a glutamic acid at position 186. However, approximately 10 years post emergence this amino acid changed for a lysine, which abolished NS1 interaction with CPSF30 and the subsequent block of general gene expression (Chapter 4). These findings were in accordance with previous reports from the literature describing a link between viral adaption to certain host species and the loss of CPSF30 binding (Hale et al., 2010, Brown et al., 2001, Hossain et al., 2008). Interestingly, NS1 residue 186 is a direct neighbour of the highly conserved tryptophan 187, which is known to regulate NS1 effector domain (ED) dimerization (Hale et al., 2008, Kerry et al., 2011). This dimerization has been shown to reinforce NS1 anti-IFN activity by strengthening dsRNA binding (Aramini et al., 2011). More importantly NS1 ED dimerization was shown to be incompatible with CPSF30 binding (Kerry et al., 2011). Thus, it is possible that the loss of

CPSF30 binding of EIV NS1 ten years post emergence, has allowed the protein to interact with new binding partners or unmasked or reinforced important functions that were selectively advantageous for EIV, such as anti-IFN activities. In accordance with this hypothesis, transfection experiments also revealed that the repression exerted by NS1 over the IFN-β induction remained strong throughout evolution, which suggested that maintaining control over one of the main anti-viral defence of the host was selectively advantageous for the virus (Chapter 4).

Twenty years post emergence EIV NS1 was truncated of eleven-amino acids at its C-terminus, due to a non-sense mutation at codon 220. Although the effect of this truncation on NS1 function was not as marked as the E186K substitution, it seemed to affect NS1 control of general gene expression and ISG induction in transfection experiments.

In the future, it would be interesting to dissect in more details the contribution of NS1 residue 186 and C-terminal tail in the protein anti-IFN action, as well as their potential contribution to other NS1 function.

The impact of NS1 evolutionary markers, E186K and C-terminal truncation, on viral infection phenotype were then evaluated. To this end, two eight-plasmid reverse genetic systems, A/equine/Uruguay/1/1963 and A/equine/Ohio/1/2003, were used. U/63 is believed to be at the origin of the H3N8 EIV lineage and it naturally possesses a full-length NS1 protein harbouring E186. O/03 is the representative of Florida Clade 1 viruses currently circulating in the equine population, and expresses naturally a C-terminally truncated NS1 protein harbouring K186. By site-directed mutagenesis (Chapter 5), mutant versions of these two viruses for NS1 residue 186 and C-terminus were engineered. Their ability to grow and spread infection in mammalian cells were then evaluated and compared with that of their respective wild-type viruses. Surprisingly, NS1 evolutionary markers appeared to affect viral infection phenotype in a virus-context and cell type-dependent manner. For instance, the viral fitness of the contemporary virus (O/03) seemed to strongly depend on the maintenance of both NS1 evolutionary markers, K186 and C-terminal truncation. Indeed, a short NS1 with K186 was essential for O/03 replication in the presence of type I IFN, as well as to limit the establishment of an antiviral state, and delay apoptosis (Chapter 5). These characteristics are likely to be significant fitness traits, as they not only render an important arm of the host antiviral response ineffective, but they also maintain the cellular resources available for longer. Unexpectedly, NS1 K186E substitution seemed to have a high fitness cost for the contemporary virus in equine cells, particularly when introduced in the context of a short NS1 protein (Chapter 5). Indeed, the O/03-K186E mutant virus did not grow in equine cells, and did not induce protein shutdown or apoptosis, and no ISG or viral protein

could be detected during infection. Interestingly, the O/03 double mutant virus containing also the K186E substitution was not as strongly attenuated. Indeed, this virus could replicate in equine cells, produce its own viral proteins, and shutdown cellular protein production at early times post infection. Several hypotheses can be formulated to explain these discrepancies: (1) potential new functions provided by the extension of the NS1 C-terminal tail, e.g. PABPII binding or interaction with PDZ-containing proteins (Obenauer et al., 2006), could have partially compensated for the loss of other functions following the restitution of CPSF30 binding; (2) the lack of protein shutoff by O/03-K186E single mutant virus could simply be due to the inability of this virus to express its own proteins, including NS1; (3) finally, to be able to shutdown protein production in equine cells a full length NS1 protein with a glutamic acid at position 186 could be necessary. Further work would be needed to test these hypotheses.

In addition, both O/03 mutant viruses expressing a full-length NS1 protein were inducing apoptosis prematurely, and even more so when associated with E186K substitution. Accordingly, amino acids 181 to 185 have previously been involved in apoptosis delay via activation of the PI3K/Akt-pathway (Ehrhardt et al., 2006, Ehrhardt et al., 2007, Hale et al., 2006, Shin et al., 2007). However, the introduced mutations could also have affected the potential control of PKR activation (Fujimoto et al., 1998, Takizawa et al., 1996, Takizawa et al., 1995, Wada et al., 1995). Indeed, following dsRNA sensing, PKR is activated and starts a cascade of events leading to shutoff of protein production (Bergmann et al., 2000, Hatada et al., 1999) and induction of apoptosis (Takizawa et al., 1996, Van Campen et al., 1989). It would be worth testing if O/03 WT NS1 and mutant can equally interfere with PKR function or activate the PI3K pathway.

For U/63, although the effect of NS1 C-terminal truncation was negligible, the loss of CPSF30 binding (via E186K substitution) reduced significantly growth kinetics at early times post infection. Since the capacity to rapidly produce a new progeny is likely to be important for viral emergence, the glutamic acid at position 186, and subsequent ability to interact with CPSF30 must have played an important role in EIV emergence.

Interestingly, a similar evolution pattern to EIV NS1 has been observed for the North American ‘classical’ swine H1N1 lineage, whose NS segment is of avian origin. This virus maintained a full length NS1 protein until the mid-1960s, before introducing of a stop codon at position 220, as for EIV NS1, which resulted in an eleven-amino acids C-terminal truncation of the protein. These changes have subsequently been retained in the ‘classical’ swine H1N1 lineage until the present day (Hale et al., 2010).

Thus, the loss of CPSF30 binding and C-terminal truncation of NS1 seem to be a common evolutionary trait between swine and equine Influenza A viruses. Interestingly, human influenza A viruses seem somewhat different, and preferentially select for NS1 proteins binding CPSF30. For example, the NS1 protein of H5N1 viruses presented a defect in inhibiting general gene expression when the transmission occurred from birds to humans in 1997, however the viruses isolated since 1998 have gained this NS1 function (Twu et al., 2007, Clark et al., 2017). It would be interesting to compare the advantages and downsides of losing the CPSF30 binding and or the C-terminal tail in different hosts species.

To evaluate the impact of NS1 evolutionary markers on virus-host interaction the early gene expression profile of equine cells infected with wild-type and NS1 mutant viruses were analysed (Chapter 6). NS1 residue 186 and C-terminal tail were shown to affect host and viral gene expression in a virus-context dependent manner. The concomitant introduction of K186E substitution and C-terminal extension in O/03 NS1 led to an increased expression of HA, NA and NS viral genes, and decreased NP gene expression. Interestingly, previous reports from the literature showed that an increased expression of HA, NA and NS were associated with an increased replication efficiency early post infection (Park et al., 2015). Feature that was also observed for O/03-K186E-230 (Chapter 5). Although it was beyond the scope of this work, it would be of great interest to analyse in details the mechanisms behind this phenotype.

To identify possible biological interactions of cellular genes differentially expressed upon infection and identify important functional networks affected by EIV and modulated by NS1 evolutionary markers, the Ingenuity pathway analysis tool was used (Chapter 6). It was observed that the wild-type O/03 virus was able to activate a large number of cellular pathways, and notably to up-regulate cell survival pathways, such as the PI3K/AKT pathway, while down-regulating pathways associated with cell cycle arrest and cell death. More importantly, a short NS1 protein with K186 was important for these functions, as none of the O/03 mutants were able to do so.

In addition, all viruses apart from the wild-type O/03 virus were unable to control several key anti-viral players, such as IFN, IRF, or RIG-I pathways (Chapter 6), which was consistent with their attenuated phenotype in interferon competent equine cells (Chapter 5). This further emphasises the importance of NS1 evolutionary markers in EIV adaptation to horses, but also suggests that NS1 in not sufficient to improve viral fitness of an emergent virus.

Finally, some inflammatory pathways were downregulated by the wild-type O/03 virus, such as IL-17A, IL-6 or TNFR1 signalling, while others were activated, such as LPS/IL-1, FcεRI, or IL-8 signalling. More importantly, the U/63 double mutant virus positively regulated the FcεRI signalling pathway as well, which has previously been associated with asthma pathogenesis (Lloyd and Marsland, 2017). This raises questions of the impact of exposure to Influenza viruses on exacerbation of chronic lung diseases.

Altogether these data suggest that the emergent EIV, which lacked the ability to control efficiently the host antiviral response, relied on a strong shutdown of general gene expression via an NS1-CPSF30-dependent mechanism. In doing so, the un-adapted virus created a short window during which it could take control of host gene expression, and produce and transmit a new progeny, before being outcompeted by the host immune response. But, this approach had strong negative impacts on host cell homeostasis, and was associated with to a premature induction of apoptosis. However, as the virus evolved in the new host, genetic mutations arose, and those that improved control of the host anti-viral responses while promoting cell survival were preferentially selected. This allowed the maintenance of the cellular resources for longer by delaying apoptosis induction, and overall increased replication efficiency (Figure 7.1).

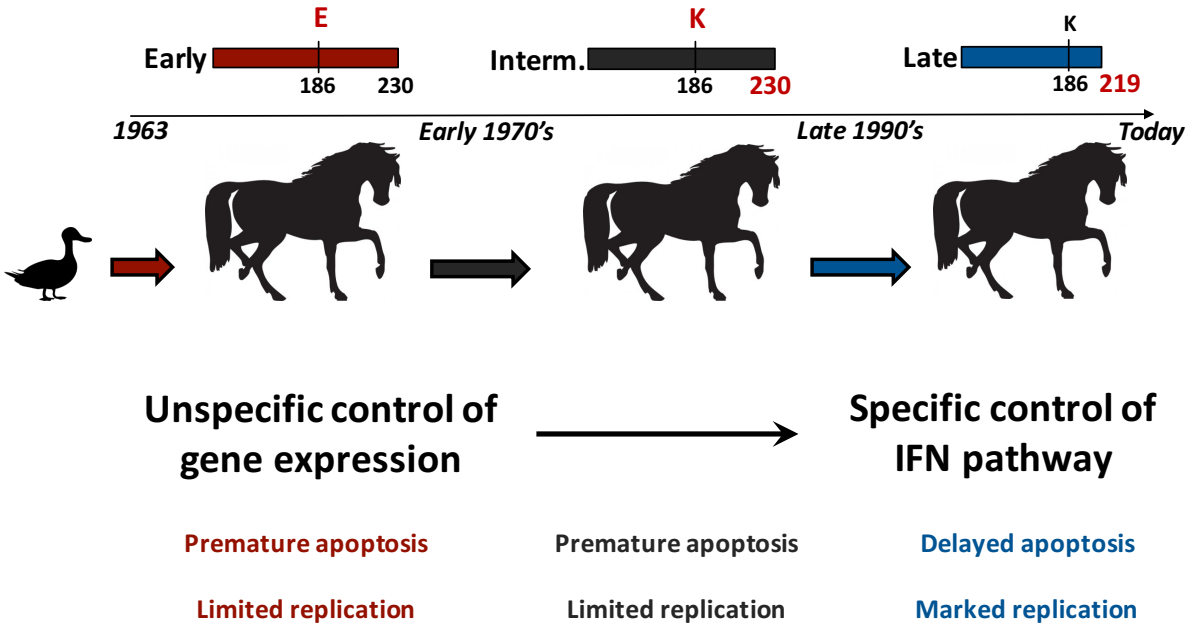
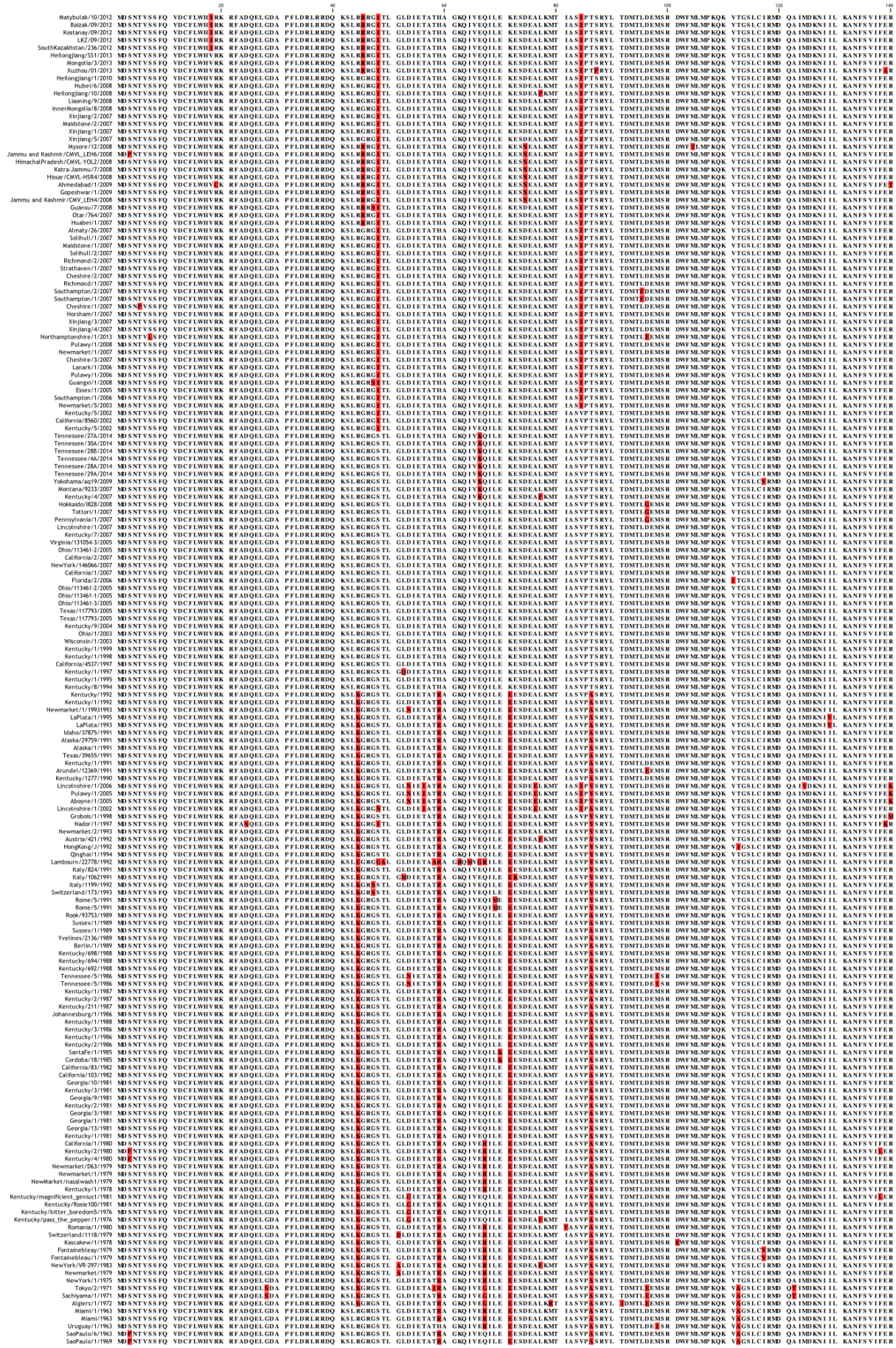


Figure 7.1: Summary of the functional evolution of the EIV NS1 protein from viral emergence to date

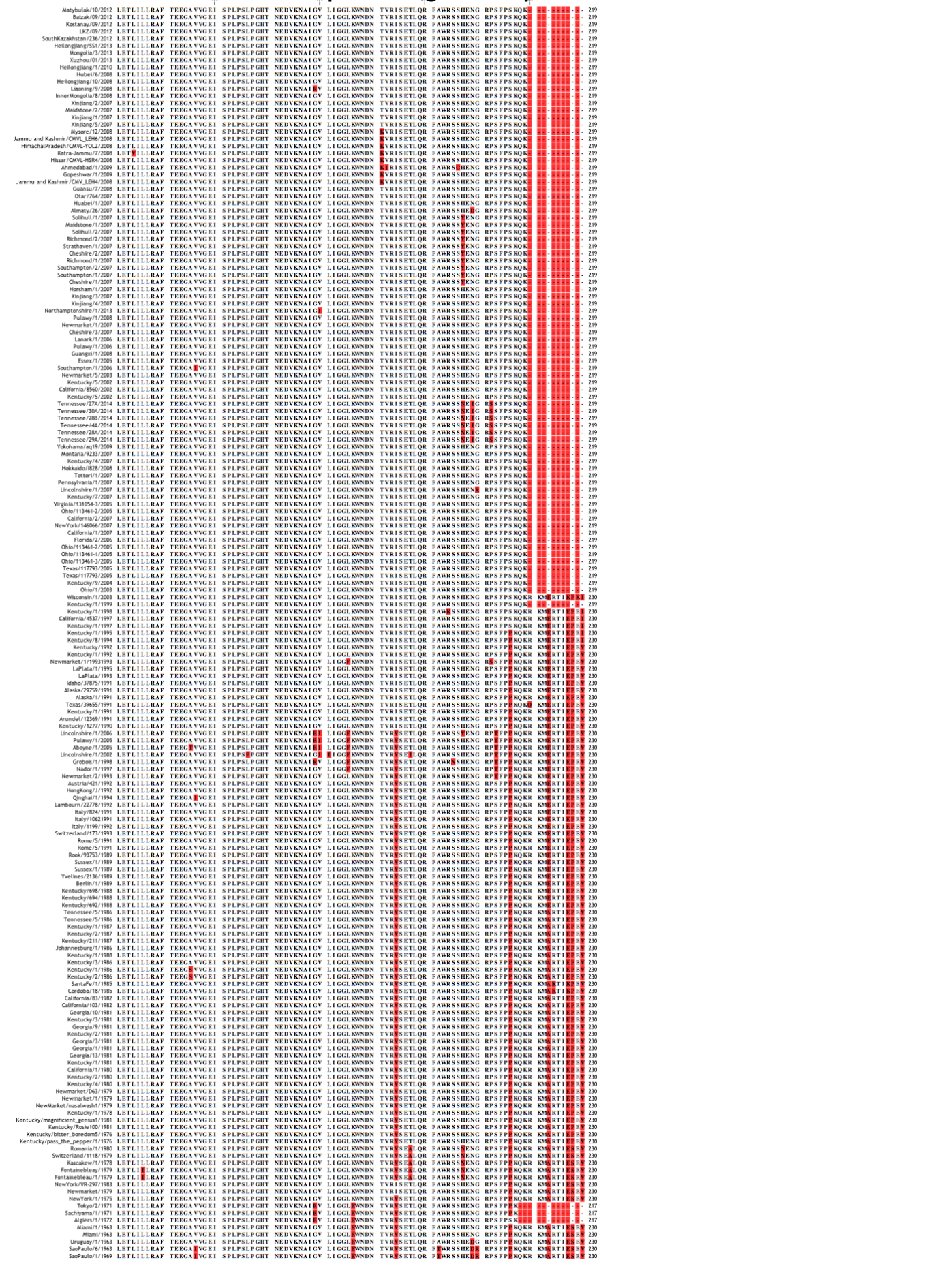
A schematic representation of the molecular evolution of the NS1 protein of the H3N8 EIV during 50 years of circulation of the virus in the equine population is shown. Markers of evolution of the NS1 protein are indicated in red, and the consequences of these changes on virus-host interaction and consequences on EIV infection phenotype are indicated.

APPENDICES

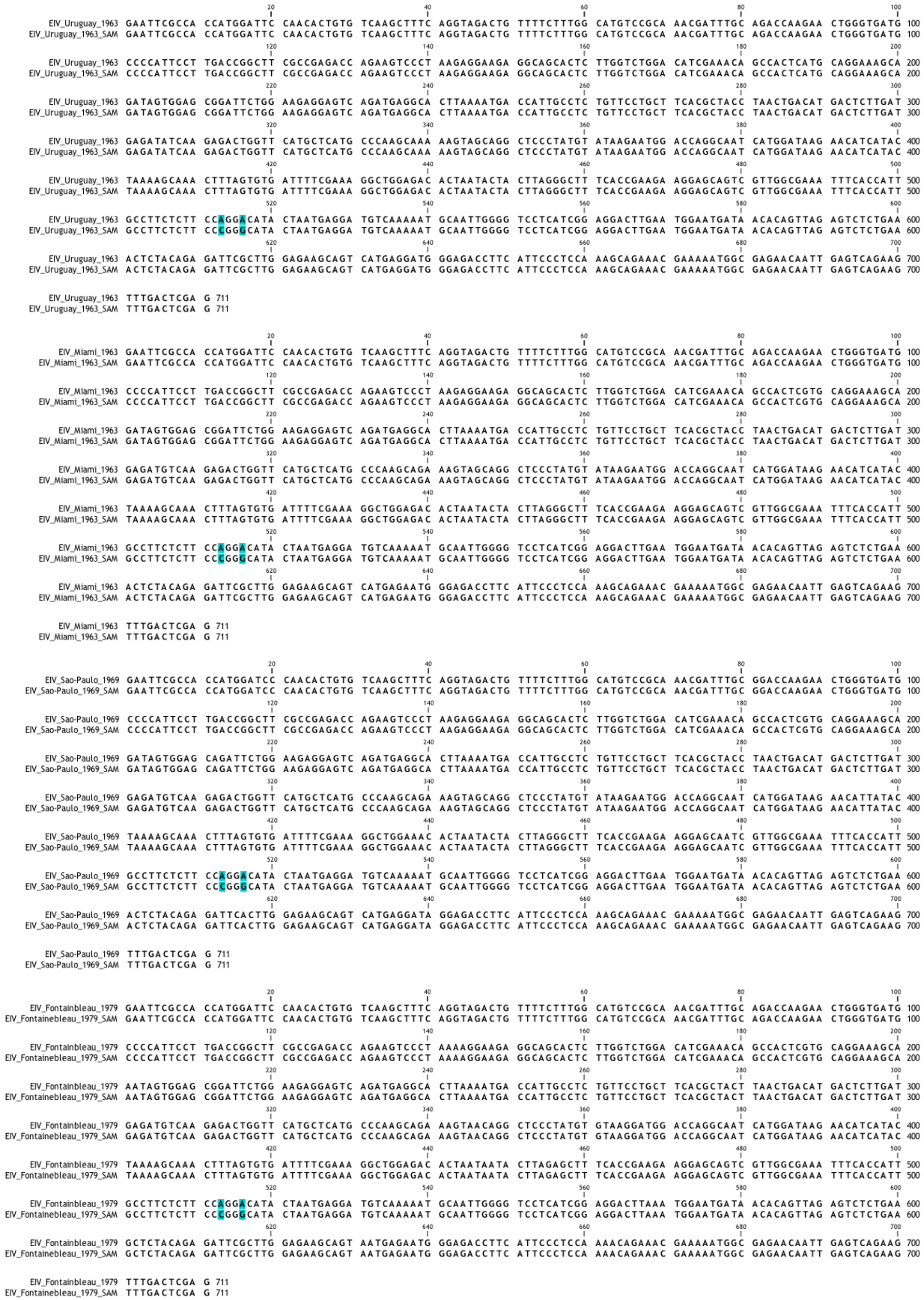
APPENDIX 1: NS1 amino acid sequence alignment - part 1



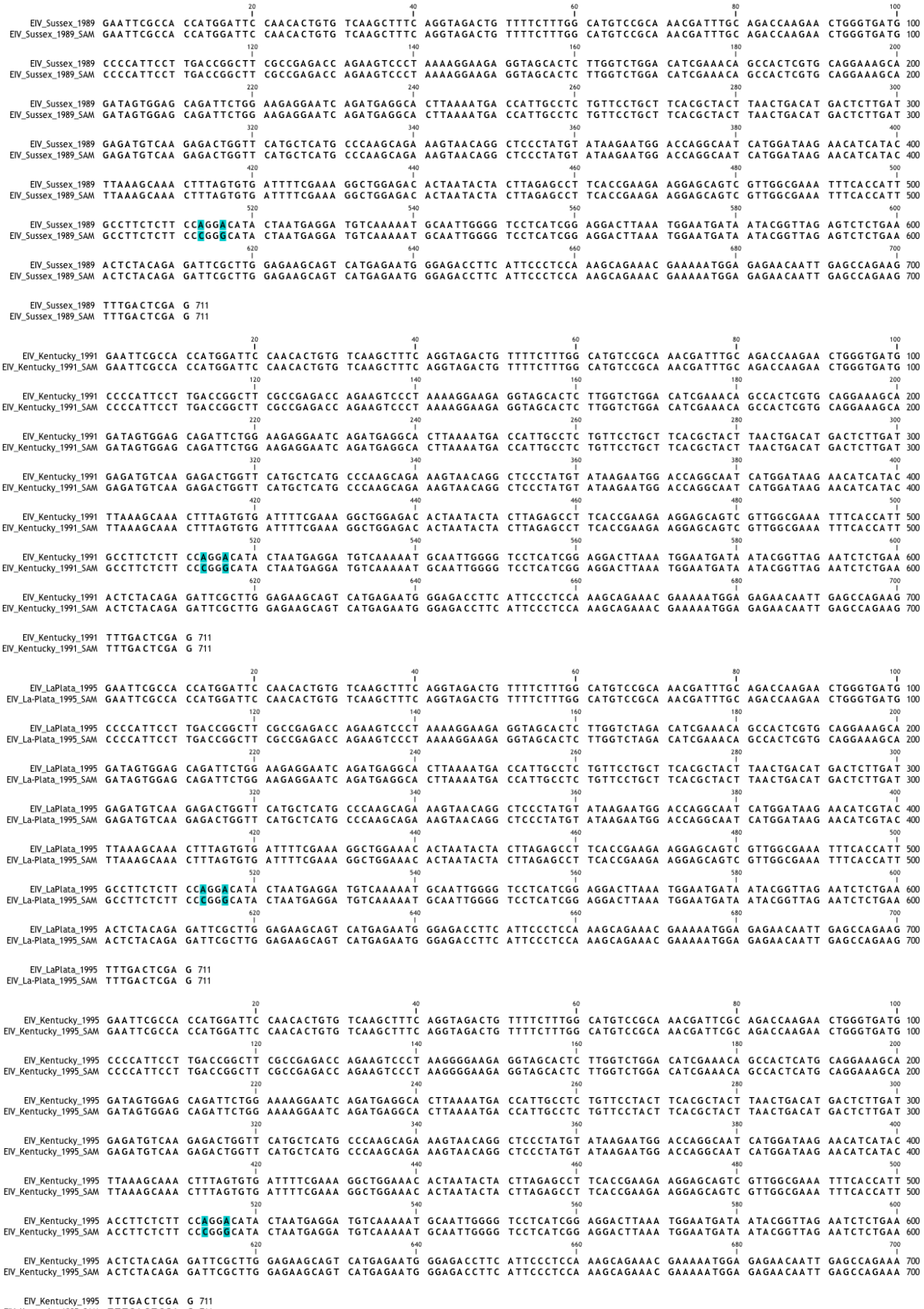
APPENDIX 2: NS1 amino acid sequence alignment - part 2



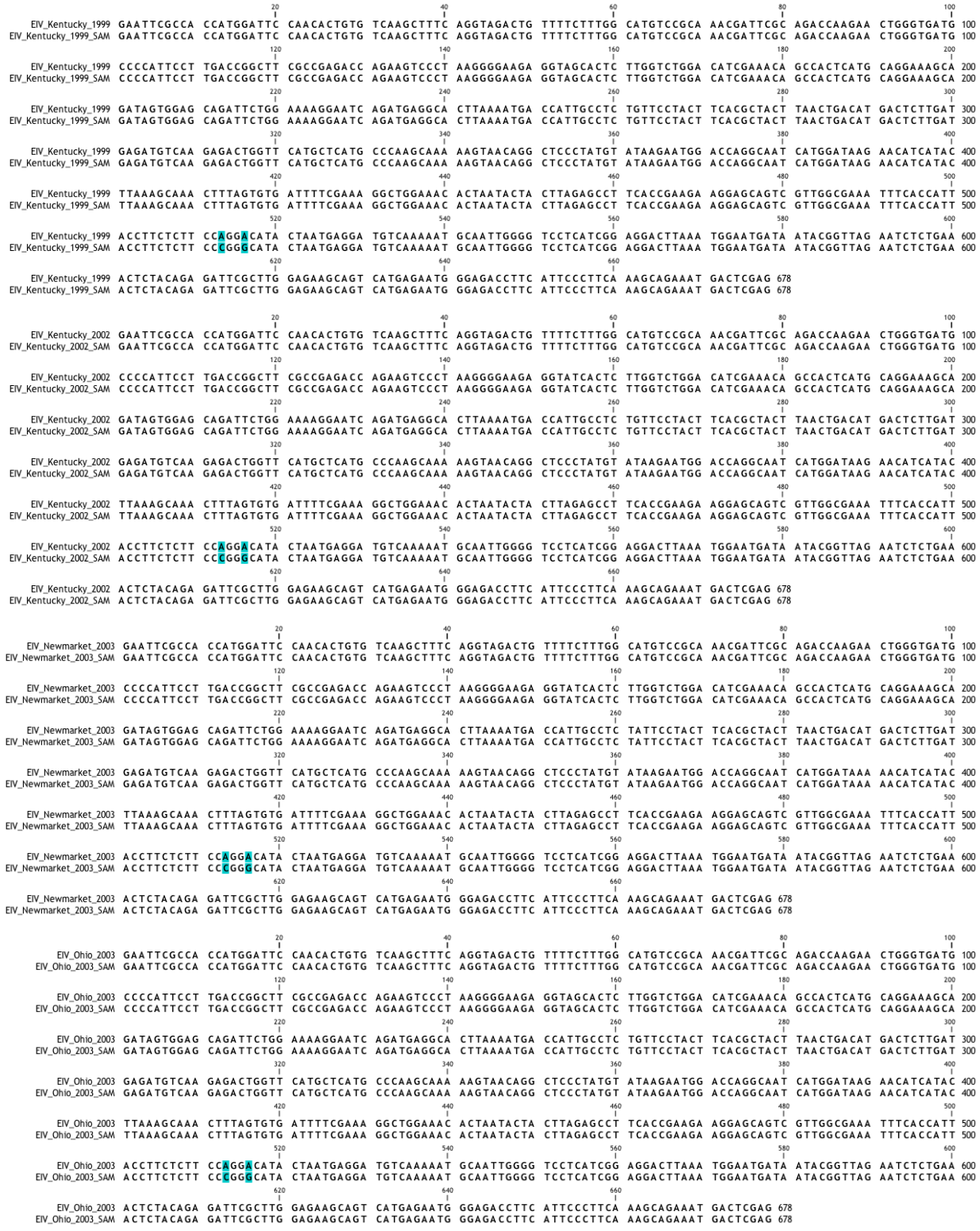
APPENDIX 3: Alignment of NS1 WT and SAM pCAGGS constructs – Part 1



APPENDIX 4: Alignment of NS1 WT and SAM pCAGGS constructs – Part 2



APPENDIX 5: Alignment of NS1 WT and SAM pCAGGS constructs – Part 3



APPENDIX 6: Alignment of NS1 WT and SAM pCAGGS constructs – Part 4

APPENDIX 7: Sequences of 13 representative EIVs – Part 1

- **A/equine/Uruguay/1/1963**

NS segment (CY032425): 890 bp

```
1    AGCAAAAGCA GGGTGACAAA AACATAATGG ATTCCAACAC TGTGTCAAGC TTTCAGGTAG
61   ACTGTTTTCT TTGGCATGTC CGCAAACGAT TTGCAGACCA AGAACTGGGT GATGCCCCAT
121  TCCTTGACCG GCTTCGCCGA GACCAGAAGT CCCTAAGAGG AAGAGGCAGC ACTCTTGGTC
181  TGGACATCGA AACAGCCACT CATGCAGGAA AGCAGATAGT GGAGCGGATT CTGGAAGAGG
241  AGTCAGATGA GGCACTTAAA ATGACCATTG CCTCTGTTCC TGCTTCACGC TACCTAACTG
301  ACATGACTCT TGATGAGATA TCAAGAGACT GGTTCATGCT CATGCCCAAG CAAAAAGTAG
361  CAGGCTCCCT ATGTATAAGA ATGGACCAGG CAATCATGGA TAAGAACATC ATACTAAAAG
421  CAAACTTTAG TGTGATTTTC GAAAGGCTGG AGACACTAAT ACTACTTAGG GCTTTCACCG
481  AAGAAGGAGC AGTCGTTGGC GAAATTTCAC CATTGCCTTC TCTTCCAGGA CATACTAATG
541  AGGATGTCAA AAATGCAATT GGGGTCCTCA TCGGAGGACT TGAATGGAAT GATAACACAG
601  TTAGAGTCTC TGAAACTCTA CAGAGATTCG CTTGGAGAAG CAGTCATGAG GATGGGAGAC
661  CTTCATTCCC TCCAAAGCAG AAACGAAAAA TGGCGAGAAC AATTGAGTCA GAAGTTTGAA
721  GAAATAAGGT GGTTGATTGA AGAAGTGCGA CATAGACTGA AAATTACAGA AAATAGTTTT
781  GAACAAATAA CATTTATGCA AGCCTTACAA CTATTGCTTG AAGTGGAACA AGAGATAAGA
841  ACTTTCTCGT TTCAGCTTAT TTAATGATAA AAAACACCCT TGTTTCTACT
```

NS1 CDS (ACD85423): 693 bp

```
1    ATGGATTCCA ACACTGTGTC AAGCTTTCAG GTAGACTGTT TTCTTTGGCA TGTCCGCAAA
61   CGATTTGCAG ACCAAGAACT GGGTGATGCC CCATTCCTTG ACCGGCTTCG CCGAGACCAG
121  AAGTCCCTAA GAGGAAGAGG CAGCACTCTT GGTCTGGACA TCGAAACAGC CACTCATGCA
181  GGAAAGCAGA TAGTGGAGCG GATTCTGGAA GAGGAGTCAG ATGAGGCACT TAAAATGACC
241  ATTGCCTCTG TTCCTGCTTC ACGCTACCTA ACTGACATGA CTCTTGATGA GATATCAAGA
301  GACTGGTTCA TGCTCATGCC CAAGCAAAAA GTAGCAGGCT CCCTATGTAT AAGAATGGAC
361  CAGGCAATCA TGGATAAGAA CATCATACTA AAAGCAAACT TTAGTGTGAT TTTCGAAAGG
421  CTGGAGACAC TAATACTACT TAGGGCTTTC ACCGAAGAAG GAGCAGTCGT TGGCGAAATT
481  TCACCATTGC CTTCTCTTCC AGGACATACT AATGAGGATG TCAAAAATGC AATTGGGGTC
541  CTCATCGGAG GACTTGAATG GAATGATAAC ACAGTTAGAG TCTCTGAAAC TCTACAGAGA
601  TTCGCTTGGA GAAGCAGTCA TGAGGATGGG AGACCTTCAT TCCCTCCAAA GCAGAAACGA
661  AAAATGGCGA GAACAATTGA GTCAGAAGTT TGA
```

NS1 amino acid sequence (ACD85423): 230 aa

```
1    MDSNTVSSFQ VDCFLWHVRK RFADQELGDA PFLDRLRRDQ KSLRGRGSTL GLDIETATHA
61   GKQIVERILE EESDEALKMT IASVPASRYL TDMTLDEISR DWFMLMPKQK VAGSLCIRMD
121  QAIMDKNIIL KANFSVIFER LETLILLRAF TEEGAVVGEI SPLPSLPGHT NEDVKNAIGV
181  LIGGLEWNDN TVRVSETLQR FAWRSSHEDG RPSFPPKQKR KMARTIESEV
```

APPENDIX 8: Sequences of 13 representative EIVs – Part 2

- **A/equine/Miami/1/1963**

NS segment (CY028840): 890 bp

```
1    AGCAAAAGCA GGGTGACAAA AACATAATGG ATTCCAACAC TGTGTCAAGC TTTCAGGTAG
61   ACTGTTTTCT TTGGCATGTC CGCAAACGAT TTGCAGACCA AGAACTGGGT GATGCCCCAT
121  TCCTTGACCG GCTTCGCCGA GACCAGAAGT CCCTAAGAGG AAGAGGCAGC ACTCTTGGTC
181  TGGACATCGA AACAGCCACT CGTGCAGGAA AGCAGATAGT GGAGCGGATT CTGGAAGAGG
241  AGTCAGATGA GGCACTTAAA ATGACCATTG CCTCTGTTCC TGCTTCACGC TACCTAACTG
301  ACATGACTCT TGATGAGATG TCAAGAGACT GGTTCATGCT CATGCCCAAG CAGAAAGTAG
361  CAGGCTCCCT ATGTATAAGA ATGGACCAGG CAATCATGGA TAAGAACATC ATACTAAAAG
421  CAAACTTTAG TGTGATTTTC GAAAGGCTGG AGACACTAAT ACTACTTAGG GCTTTCACCG
481  AAGAAGGAGC AGTCGTTGGC GAAATTTCAC CATTGCCTTC TCTTCCAGGA CATACTAATG
541  AGGATGTCAA AAATGCAATT GGGGTCCTCA TCGGAGGACT TGAATGGAAT GATAACACAG
601  TTAGAGTCTC TGAAACTCTA CAGAGATTCG CTTGGAGAAG CAGTCATGAG AATGGGAGAC
661  CTTCATTCCC TCCAAAGCAG AAACGAAAAA TGGCGAGAAC AATTGAGTCA GAAGTTTGAA
721  GAAATAAGGT GGTTGATTGA AGAAGTGCGA CATAGACTGA AAATTACAGA AAATAGTTTT
781  GAACAAATAA CATTTATGCA AGCCTTACAA CTATTGCTTG AAGTGGAACA AGAGATAAGA
841  ACTTTCTCGT TTCAGCTTAT TTAATGATAA AAAACACCCT TGTTTCTACT
```

NS1 CDS (ABY81497): 693 bp

```
1    ATGGATTCCA ACACTGTGTC AAGCTTTCAG GTAGACTGTT TTCTTTGGCA TGTCCGCAAA
61   CGATTTGCAG ACCAAGAACT GGGTGATGCC CCATTCCTTG ACCGGCTTCG CCGAGACCAG
121  AAGTCCCTAA GAGGAAGAGG CAGCACTCTT GGTCTGGACA TCGAAACAGC CACTCGTGCA
181  GGAAAGCAGA TAGTGGAGCG GATTCTGGAA GAGGAGTCAG ATGAGGCACT TAAAATGACC
241  ATTGCCTCTG TTCCTGCTTC ACGCTACCTA ACTGACATGA CTCTTGATGA GATGTCAAGA
301  GACTGGTTCA TGCTCATGCC CAAGCAGAAA GTAGCAGGCT CCCTATGTAT AAGAATGGAC
361  CAGGCAATCA TGGATAAGAA CATCATACTA AAAGCAAACT TTAGTGTGAT TTTCGAAAGG
421  CTGGAGACAC TAATACTACT TAGGGCTTTC ACCGAAGAAG GAGCAGTCGT TGGCGAAATT
481  TCACCATTGC CTTCTCTTCC AGGACATACT AATGAGGATG TCAAAAATGC AATTGGGGTC
541  CTCATCGGAG GACTTGAATG GAATGATAAC ACAGTTAGAG TCTCTGAAAC TCTACAGAGA
601  TTCGCTTGGA GAAGCAGTCA TGAGAATGGG AGACCTTCAT TCCCTCCAAA GCAGAAACGA
661  AAAATGGCGA GAACAATTGA GTCAGAAGTT TGA
```

NS1 amino acid sequence (ABY81497): 230 aa

```
1    MDSNTVSSFQ VDCFLWHVRK RFADQELGDA PFLDRLRRDQ KSLRGRGSTL GLDIETATRA
61   GKQIVERILE EESDEALKMT IASVPASRYL TDMTLDEMSR DWFMLMPKQK VAGSLCIRMD
121  QAIMDKNIIL KANFSVIFER LETLILLRAF TEEGAVVGEI SPLPSLPGHT NEDVKNAIGV
181  LIGGLEWNDN TVRVSETLQR FAWRSSHENG RPSFPPKQKR KMARTIESEV
```

APPENDIX 9: Sequences of 13 representative EIVs – Part 3

- **A/equine/SaoPaulo/1/1969**

NS segment (CY032401): 890 bp

```
1    AGCAAAAGCA GGGTGACAAA AACATAATGG ATCCCAACAC TGTGTCAAGC TTTCAGGTAG
61   ACTGTTTTCT TTGGCATGTC CGCAAACGAT TTGCGGACCA AGAACTGGGT GATGCCCCAT
121  TCCTTGACCG GCTTCGCCGA GACCAGAAGT CCCTAAGAGG AAGAGGCAGC ACTCTTGGTC
181  TGGACATCGA AACAGCCACT CGTGCAGGAA AGCAGATAGT GGAGCAGATT CTGGAAGAGG
241  AGTCAGATGA GGCACTTAAA ATGACCATTG CCTCTGTTCC TGCTTCACGC TACCTAACTG
301  ACATGACTCT TGATGAGATG TCAAGAGACT GGTTCATGCT CATGCCCAAG CAGAAAGTAG
361  CAGGCTCCCT ATGTATAAGA ATGGACCAGG CAATCATGGA TAAGAACATT ATACTAAAAG
421  CAAACTTTAG TGTGATTTTC GAAAGGCTGG AAACACTAAT ACTACTTAGG GCTTTCACCG
481  AAGAAGGAGC AATCGTTGGC GAAATTTCAC CATTGCCTTC TCTTCCAGGA CATACTAATG
541  AGGATGTCAA AAATGCAATT GGGGTCCTCA TCGGAGGACT TGAATGGAAT GATAACACAG
601  TTAGAGTCTC TGAAACTCTA CAGAGATTCA CTTGGAGAAG CAGTCATGAG GATAGGAGAC
661  CTTCATTCCC TCCAAAGCAG AAACGAAAAA TGGCGAGAAC AATTGAGTCA GAAGTTTGAA
721  GAAATAAGGT GGTTGATTGA AGAAGTGCGA CATAGACTGA AAATTACAGA AAATAGTTTT
781  GAACAAATAA CATTTATGCA AGCCTTACAA CTATTGCTTG AAGTGGAACA AGAGATAAGA
841  ACTTTCTCGT TTCAGCTTAT TTAATGATAA AAAACACCCT TGTTTCTACT
```

NS1 CDS (ACD85390): 693 bp

```
1    ATGGATCCCA ACACTGTGTC AAGCTTTCAG GTAGACTGTT TTCTTTGGCA TGTCCGCAAA
61   CGATTTGCGG ACCAAGAACT GGGTGATGCC CCATTCCTTG ACCGGCTTCG CCGAGACCAG
121  AAGTCCCTAA GAGGAAGAGG CAGCACTCTT GGTCTGGACA TCGAAACAGC CACTCGTGCA
181  GGAAAGCAGA TAGTGGAGCA GATTCTGGAA GAGGAGTCAG ATGAGGCACT TAAAATGACC
241  ATTGCCTCTG TTCCTGCTTC ACGCTACCTA ACTGACATGA CTCTTGATGA GATGTCAAGA
301  GACTGGTTCA TGCTCATGCC CAAGCAGAAA GTAGCAGGCT CCCTATGTAT AAGAATGGAC
361  CAGGCAATCA TGGATAAGAA CATTATACTA AAAGCAAACT TTAGTGTGAT TTTCGAAAGG
421  CTGGAAACAC TAATACTACT TAGGGCTTTC ACCGAAGAAG GAGCAATCGT TGGCGAAATT
481  TCACCATTGC CTTCTCTTCC AGGACATACT AATGAGGATG TCAAAAATGC AATTGGGGTC
541  CTCATCGGAG GACTTGAATG GAATGATAAC ACAGTTAGAG TCTCTGAAAC TCTACAGAGA
601  TTCACTTGGA GAAGCAGTCA TGAGGATAGG AGACCTTCAT TCCCTCCAAA GCAGAAACGA
661  AAAATGGCGA GAACAATTGA GTCAGAAGTT TGA
```

NS1 amino acid sequence (ACD85390): 230 aa

```
1    MDPNTVSSFQ VDCFLWHVRK RFADQELGDA PFLDRLRRDQ KSLRGRGSTL GLDIETATRA
61   GKQIVEQILE EESDEALKMT IASVPASRYL TDMTLDEMSR DWFMLMPKQK VAGSLCIRMD
121  QAIMDKNIIL KANFSVIFER LETLILLRAF TEEGAIVGEI SPLPSLPGHT NEDVKNAIGV
181  LIGGLEWNDN TVRVSETLQR FTWRSSHEDR RPSFPPKQKR KMARTIESEV
```

APPENDIX 10: Sequences of 13 representative EIVs – Part 4

- **A/equine/Fontainebleau/1/1979**

NS segment (CY032409): 890 bp

```
1    AGCAAAAGCA GGGTGACAAA AACATAATGG ATTCCAACAC TGTGTCAAGC TTTCAGGTAG
61   ACTGTTTTCT TTGGCATGTC CGCAAACGAT TTGCAGACCA AGAACTGGGT GATGCCCCAT
121  TCCTTGACCG GCTTCGCCGA GACCAGAAGT CCCTAAAAGG AAGAGGCAGC ACTCTTGGTC
181  TGGACATCGA AACAGCCACT CGTGCAGGAA AGCAAATAGT GGAGCGGATT CTGGAAGAGG
241  AGTCAGATGA GGCACTTAAA ATGACCATTG CCTCTGTTCC TGCTTCACGC TACTTAACTG
301  ACATGACTCT TGATGAGATG TCAAGAGACT GGTTCATGCT CATGCCCAAG CAGAAAGTAA
361  CAGGCTCCCT ATGTGTAAGG ATGGACCAGG CAATCATGGA TAAGAACATC ATACTAAAAG
421  CAAACTTTAG TGTGATTTTC GAAAGGCTGG AGACACTAAT AATACTTAGA GCTTTCACCG
481  AAGAAGGAGC AGTCGTTGGC GAAATTTCAC CATTGCCTTC TCTTCCAGGA CATACTAATG
541  AGGATGTCAA AAATGCAATT GGGGTCCTCA TCGGAGGACT TAAATGGAAT GATAACACAG
601  TTAGAGTCTC TGAAGCTCTA CAGAGATTCG CTTGGAGAAG CAGTAATGAG AATGGGAGAC
661  CTTCATTCCC TCCAAAACAG AAACGAAAAA TGGCGAGAAC AATTGAGTCA GAAGTTTGAA
721  GAAATAAGAT GGTTGATTGA AGAAGTGCGA CATAGATTGA AAAATACAGA AAATAGTTTT
781  GAACAAATCA CATTTATGCA AGCCTTACAA CTATTGCTTG AAGTAGAACA AGAGATAAGA
841  ACTTTCTCGT TTCAGCTTAT TTAATGATAA AAAACACCCT TGTTTCTACT
```

NS1 CDS (ACD85401): 693 bp

```
1    ATGGATTCCA ACACTGTGTC AAGCTTTCAG GTAGACTGTT TTCTTTGGCA TGTCCGCAAA
61   CGATTTGCAG ACCAAGAACT GGGTGATGCC CCATTCCTTG ACCGGCTTCG CCGAGACCAG
121  AAGTCCCTAA AAGGAAGAGG CAGCACTCTT GGTCTGGACA TCGAAACAGC CACTCGTGCA
181  GGAAAGCAAA TAGTGGAGCG GATTCTGGAA GAGGAGTCAG ATGAGGCACT TAAAATGACC
241  ATTGCCTCTG TTCCTGCTTC ACGCTACTTA ACTGACATGA CTCTTGATGA GATGTCAAGA
301  GACTGGTTCA TGCTCATGCC CAAGCAGAAA GTAACAGGCT CCCTATGTGT AAGGATGGAC
361  CAGGCAATCA TGGATAAGAA CATCATACTA AAAGCAAACT TTAGTGTGAT TTTCGAAAGG
421  CTGGAGACAC TAATAATACT TAGAGCTTTC ACCGAAGAAG GAGCAGTCGT TGGCGAAATT
481  TCACCATTGC CTTCTCTTCC AGGACATACT AATGAGGATG TCAAAAATGC AATTGGGGTC
541  CTCATCGGAG GACTTAAATG GAATGATAAC ACAGTTAGAG TCTCTGAAGC TCTACAGAGA
601  TTCGCTTGGA GAAGCAGTAA TGAGAATGGG AGACCTTCAT TCCCTCCAAA ACAGAAACGA
661  AAAATGGCGA GAACAATTGA GTCAGAAGTT TGA
```

NS1 amino acid sequence (ACD85401): 230 aa

```
1    MDSNTVSSFQ VDCFLWHVRK RFADQELGDA PFLDRLRRDQ KSLKGRGSTL GLDIETATRA
61   GKQIVERILE EESDEALKMT IASVPASRYL TDMTLDEMSR DWFMLMPKQK VTGSLCVRMD
121  QAIMDKNIIL KANFSVIFER LETLIILRAF TEEGAVVGEI SPLPSLPGHT NEDVKNAIGV
181  LIGGLKWNDN TVRVSEALQR FAWRSSNENG RPSFPPKQKR KMARTIESEV
```

APPENDIX 11: Sequences of 13 representative EIVs – Part 5

- **A/equine/Sussex/1/1989**

NS segment (CY032321): 890 bp

```
1    AGCAAAAGCA GGGTGACAAA AACATAATGG ATTCCAACAC TGTGTCAAGC TTTCAGGTAG
61   ACTGTTTTCT TTGGCATGTC CGCAAACGAT TTGCAGACCA AGAACTGGGT GATGCCCCAT
121  TCCTTGACCG GCTTCGCCGA GACCAGAAGT CCCTAAAAGG AAGAGGTAGC ACTCTTGGTC
181  TGGACATCGA AACAGCCACT CGTGCAGGAA AGCAGATAGT GGAGCAGATT CTGGAAGAGG
241  AATCAGATGA GGCACTTAAA ATGACCATTG CCTCTGTTCC TGCTTCACGC TACTTAACTG
301  ACATGACTCT TGATGAGATG TCAAGAGACT GGTTCATGCT CATGCCCAAG CAGAAAGTAA
361  CAGGCTCCCT ATGTATAAGA ATGGACCAAG CAATCATGGA TAAGAACATC ATACTTAAAG
421  CAAACTTTAG TGTGATTTTC GAAAGGCTGG AGACACTAAT ACTACTTAGA GCCTTCACCG
481  AAGAAGGAGC AGTCGTTGGC GAAATTTCAC CATTGCCTTC TCTTCCAGGA CATACTAATG
541  AGGATGTCAA AAATGCAATT GGGGTCCTCA TCGGAGGACT TAAATGGAAT GATAATACGG
601  TTAGAGTCTC TGAAACTCTA CAGAGATTCG CTTGGAGAAG CAGTCATGAG AATGGGAGAC
661  CTTCATTCCC TCCAAAGCAG AAACGAAAAA TGGAGAGAAC AATTGAGCCA GAAGTTTGAA
721  GAAATAAGAT GGTTGATTGA AGAAGTGCGA CATAGATTGA AAAATACAGA AAATAGTTTT
781  GAACAAATAA CATTTATGCA AGCCTTACAA CTATTGCTTG AAGTAGAACA AGAGATAAGA
841  ACTTTCTCGT TTCAGCTTAT TTAATGATAA AAAACACCCT TGTTTCTACT
```

NS1 CDS (ACD97430): 693 bp

```
1    ATGGATTCCA ACACTGTGTC AAGCTTTCAG GTAGACTGTT TTCTTTGGCA TGTCCGCAAA
61   CGATTTGCAG ACCAAGAACT GGGTGATGCC CCATTCCTTG ACCGGCTTCG CCGAGACCAG
121  AAGTCCCTAA AAGGAAGAGG TAGCACTCTT GGTCTGGACA TCGAAACAGC CACTCGTGCA
181  GGAAAGCAGA TAGTGGAGCA GATTCTGGAA GAGGAATCAG ATGAGGCACT TAAAATGACC
241  ATTGCCTCTG TTCCTGCTTC ACGCTACTTA ACTGACATGA CTCTTGATGA GATGTCAAGA
301  GACTGGTTCA TGCTCATGCC CAAGCAGAAA GTAACAGGCT CCCTATGTAT AAGAATGGAC
361  CAAGCAATCA TGGATAAGAA CATCATACTT AAAGCAAACT TTAGTGTGAT TTTCGAAAGG
421  CTGGAGACAC TAATACTACT TAGAGCCTTC ACCGAAGAAG GAGCAGTCGT TGGCGAAATT
481  TCACCATTGC CTTCTCTTCC AGGACATACT AATGAGGATG TCAAAAATGC AATTGGGGTC
541  CTCATCGGAG GACTTAAATG GAATGATAAT ACGGTTAGAG TCTCTGAAAC TCTACAGAGA
601  TTCGCTTGGA GAAGCAGTCA TGAGAATGGG AGACCTTCAT TCCCTCCAAA GCAGAAACGA
661  AAAATGGAGA GAACAATTGA GCCAGAAGTT TGA
```

NS1 amino acid sequence (ACD97430): 230 aa

```
1    MDSNTVSSFQ VDCFLWHVRK RFADQELGDA PFLDRLRRDQ KSLKGRGSTL GLDIETATRA
61   GKQIVEQILE EESDEALKMT IASVPASRYL TDMTLDEMSR DWFMLMPKQK VTGSLCIRMD
121  QAIMDKNIIL KANFSVIFER LETLILLRAF TEEGAVVGEI SPLPSLPGHT NEDVKNAIGV
181  LIGGLKWNDN TVRVSETLQR FAWRSSHENG RPSFPPKQKR KMERTIEPEV
```

APPENDIX 12: Sequences of 13 representative EIVs – Part 6

- **A/equine/Kentucky/1/1991**

NS segment (CY030177): 890 bp

```
1    AGCAAAAGCA GGGTGACAAA AACATAATGG ATTCCAACAC TGTGTCAAGC TTTCAGGTAG
61   ACTGTTTTCT TTGGCATGTC CGCAAACGAT TTGCAGACCA AGAACTGGGT GATGCCCCAT
121  TCCTTGACCG GCTTCGCCGA GACCAGAAGT CCCTAAAAGG AAGAGGTAGC ACTCTTGGTC
181  TGGACATCGA AACAGCCACT CGTGCAGGAA AGCAGATAGT GGAGCAGATT CTGGAAGAGG
241  AATCAGATGA GGCACTTAAA ATGACCATTG CCTCTGTTCC TGCTTCACGC TACTTAACTG
301  ACATGACTCT TGATGAGATG TCAAGAGACT GGTTCATGCT CATGCCCAAG CAGAAAGTAA
361  CAGGCTCCCT ATGTATAAGA ATGGACCAGG CAATCATGGA TAAGAACATC ATACTTAAAG
421  CAAACTTTAG TGTGATTTTC GAAAGGCTGG AGACACTAAT ACTACTTAGA GCCTTCACCG
481  AAGAAGGAGC AGTCGTTGGC GAAATTTCAC CATTGCCTTC TCTTCCAGGA CATACTAATG
541  AGGATGTCAA AAATGCAATT GGGGTCCTCA TCGGAGGACT TAAATGGAAT GATAATACGG
601  TTAGAATCTC TGAAACTCTA CAGAGATTCG CTTGGAGAAG CAGTCATGAG AATGGGAGAC
661  CTTCATTCCC TCCAAAGCAG AAACGAAAAA TGGAGAGAAC AATTGAGCCA GAAGTTTGAA
721  GAAATAAGAT GGTTGATTGA AGAAGTGCGA CATAGATTGA AAAATACAGA AAATAGTTTT
781  GAACAAATAA CATTTATGCA AGCCTTACAA CTATTGCTTG AAGTAGAACA AGAGATAAGA
841  ACTTTCTCGT TTCAGCTTAT TTAATGATAA AAAACACCCT TGTTTCTACT
```

NS1 CDS (ACA24672): 693 bp

```
1    ATGGATTCCA ACACTGTGTC AAGCTTTCAG GTAGACTGTT TTCTTTGGCA TGTCCGCAAA
61   CGATTTGCAG ACCAAGAACT GGGTGATGCC CCATTCCTTG ACCGGCTTCG CCGAGACCAG
121  AAGTCCCTAA AAGGAAGAGG TAGCACTCTT GGTCTGGACA TCGAAACAGC CACTCGTGCA
181  GGAAAGCAGA TAGTGGAGCA GATTCTGGAA GAGGAATCAG ATGAGGCACT TAAAATGACC
241  ATTGCCTCTG TTCCTGCTTC ACGCTACTTA ACTGACATGA CTCTTGATGA GATGTCAAGA
301  GACTGGTTCA TGCTCATGCC CAAGCAGAAA GTAACAGGCT CCCTATGTAT AAGAATGGAC
361  CAGGCAATCA TGGATAAGAA CATCATACTT AAAGCAAACT TTAGTGTGAT TTTCGAAAGG
421  CTGGAGACAC TAATACTACT TAGAGCCTTC ACCGAAGAAG GAGCAGTCGT TGGCGAAATT
481  TCACCATTGC CTTCTCTTCC AGGACATACT AATGAGGATG TCAAAAATGC AATTGGGGTC
541  CTCATCGGAG GACTTAAATG GAATGATAAT ACGGTTAGAA TCTCTGAAAC TCTACAGAGA
601  TTCGCTTGGA GAAGCAGTCA TGAGAATGGG AGACCTTCAT TCCCTCCAAA GCAGAAACGA
661  AAAATGGAGA GAACAATTGA GCCAGAAGTT TGA
```

NS1 amino acid sequence (ACA24672): 230 aa

```
1    MDSNTVSSFQ VDCFLWHVRK RFADQELGDA PFLDRLRRDQ KSLKGRGSTL GLDIETATRA
61   GKQIVEQILE EESDEALKMT IASVPASRYL TDMTLDEMSR DWFMLMPKQK VTGSLCIRMD
121  QAIMDKNIIL KANFSVIFER LETLILLRAF TEEGAVVGEI SPLPSLPGHT NEDVKNAIGV
181  LIGGLKWNDN TVRISETLQR FAWRSSHENG RPSFPPKQKR KMERTIEPEV
```

APPENDIX 13: Sequences of 13 representative EIVs – Part 7

- **A/equine/LaPlata/1/1995**

NS segment (MF182460): 890 bp

```
1    AGCAAAAGCA GGGTGACAAA AACATAATGG ATTCCAACAC TGTGTCAAGC TTTCAGGTAG
61   ACTGTTTTCT TTGGCATGTC CGCAAACGAT TTGCAGACCA AGAACTGGGT GATGCCCCAT
121  TCCTTGACCG GCTTCGCCGA GACCAGAAGT CCCTAAAAGG AAGAGGTAGC ACTCTTGGTC
181  TAGACATCGA AACAGCCACT CGTGCAGGAA AGCAGATAGT GGAGCAGATT CTGGAAGAGG
241  AATCAGATGA GGCACTTAAA ATGACCATTG CCTCTGTTCC TGCTTCACGC TACTTAACTG
301  ACATGACTCT TGATGAGATG TCAAGAGACT GGTTCATGCT CATGCCCAAG CAGAAAGTAA
361  CAGGCTCCCT ATGTATAAGA ATGGACCAGG CAATCATGGA TAAGAACATC GTACTTAAAG
421  CAAACTTTAG TGTGATTTTC GAAAGGCTGG AAACACTAAT ACTACTTAGA GCCTTCACCG
481  AAGAAGGAGC AGTCGTTGGC GAAATTTCAC CATTGCCTTC TCTTCCAGGA CATACTAATG
541  AGGATGTCAA AAATGCAATT GGGGTCCTCA TCGGAGGACT TAAATGGAAT GATAATACGG
601  TTAGAATCTC TGAAACTCTA CAGAGATTCG CTTGGAGAAG CAGTCATGAG AATGGGAGAC
661  CTTCATTCCC TCCAAAGCAG AAACGAAAAA TGGAGAGAAC AATTGAGCCA GAAGTTTGAA
721  GAAATAAGAT GGTTGATTGA AGAAGTGCGA CATAGATTGA GAAATACAGA AAATAGTTTT
781  GAACAAATAA CATTTATGCA AGCCTTACAA CTATTGCTTG AAGTAGAACA AGAGATAAGA
841  ACTTTCTCGT TTCAGCTTAT TTAATGATAA AAAACACCCT TGTTTCTACT
```

NS1 CDS (unassigned): 693 bp

```
1    ATGGATTCCA ACACTGTGTC AAGCTTTCAG GTAGACTGTT TTCTTTGGCA TGTCCGCAAA
61   CGATTTGCAG ACCAAGAACT GGGTGATGCC CCATTCCTTG ACCGGCTTCG CCGAGACCAG
121  AAGTCCCTAA AAGGAAGAGG TAGCACTCTT GGTCTAGACA TCGAAACAGC CACTCGTGCA
181  GGAAAGCAGA TAGTGGAGCA GATTCTGGAA GAGGAATCAG ATGAGGCACT TAAAATGACC
241  ATTGCCTCTG TTCCTGCTTC ACGCTACTTA ACTGACATGA CTCTTGATGA GATGTCAAGA
301  GACTGGTTCA TGCTCATGCC CAAGCAGAAA GTAACAGGCT CCCTATGTAT AAGAATGGAC
361  CAGGCAATCA TGGATAAGAA CATCGTACTT AAAGCAAACT TTAGTGTGAT TTTCGAAAGG
421  CTGGAAACAC TAATACTACT TAGAGCCTTC ACCGAAGAAG GAGCAGTCGT TGGCGAAATT
481  TCACCATTGC CTTCTCTTCC AGGACATACT AATGAGGATG TCAAAAATGC AATTGGGGTC
541  CTCATCGGAG GACTTAAATG GAATGATAAT ACGGTTAGAA TCTCTGAAAC TCTACAGAGA
601  TTCGCTTGGA GAAGCAGTCA TGAGAATGGG AGACCTTCAT TCCCTCCAAA GCAGAAACGA
661  AAAATGGAGA GAACAATTGA GCCAGAAGTT TGA
```

NS1 amino acid sequence (unassigned): 230 aa

```
1    MDSNTVSSFQ VDCFLWHVRK RFADQELGDA PFLDRLRRDQ KSLKGRGSTL GLDIETATRA
61   GKQIVEQILE EESDEALKMT IASVPASRYL TDMTLDEMSR DWFMLMPKQK VTGSLCIRMD
121  QAIMDKNIVL KANFSVIFER LETLILLRAF TEEGAVVGEI SPLPSLPGHT NEDVKNAIGV
181  LIGGLKWNDN TVRISETLQR FAWRSSHENG RPSFPPKQKR KMERTIEPEV
```

APPENDIX 14: Sequences of 13 representative EIVs – Part 8

- **A/equine/Kentucky/1995**

NS segment (MF182451): 890 bp

```
1    AGCAAAAGCA GGGTGACAAA AACATAATGG ATTCCAACAC TGTGTCAAGC TTTCAGGTAG
61   ACTGTTTTCT TTGGCATGTC CGCAAACGAT TCGCAGACCA AGAACTGGGT GATGCCCCAT
121  TCCTTGACCG GCTTCGCCGA GACCAGAAGT CCCTAAGGGG AAGAGGTAGC ACTCTTGGTC
181  TGGACATCGA AACAGCCACT CATGCAGGAA AGCAGATAGT GGAGCAGATT CTGGAAAAGG
241  AATCAGATGA GGCACTTAAA ATGACCATTG CCTCTGTTCC TACTTCACGC TACTTAACTG
301  ACATGACTCT TGATGAGATG TCAAGAGACT GGTTCATGCT CATGCCCAAG CAGAAAGTAA
361  CAGGCTCCCT ATGTATAAGA ATGGACCAGG CAATCATGGA TAAGAACATC ATACTTAAAG
421  CAAACTTTAG TGTGATTTTC GAAAGGCTGG AAACACTAAT ACTACTTAGA GCCTTCACCG
481  AAGAAGGAGC AGTCGTTGGC GAAATTTCAC CATTACCTTC TCTTCCAGGA CATACTAATG
541  AGGATGTCAA AAATGCAATT GGGGTCCTCA TCGGAGGACT TAAATGGAAT GATAATACGG
601  TTAGAATCTC TGAAACTCTA CAGAGATTCG CTTGGAGAAG CAGTCATGAG AATGGGAGAC
661  CTTCATTCCC TCCAAAGCAG AAACGAAAAA TGGAGAGAAC AATTGAGCCA GAAATTTGAA
721  GAAATAAGAT GGTTGATTGA AGAAGTGCGA CATAGATTGA AAAATACAGA AAATAGTTTT
781  GAACAAATAA CATTTATGCA AGCCTTACAA CTATTGCTTG AAGTAGAACA AGAGATAAGA
841  ACTTTCTCGT TTCAGCTTAT TTAATGATAA AAAACACCCT TGTTTCTACT
```

NS1 CDS (unassigned): 693 bp

```
1    ATGGATTCCA ACACTGTGTC AAGCTTTCAG GTAGACTGTT TTCTTTGGCA TGTCCGCAAA
61   CGATTCGCAG ACCAAGAACT GGGTGATGCC CCATTCCTTG ACCGGCTTCG CCGAGACCAG
121  AAGTCCCTAA GGGGAAGAGG TAGCACTCTT GGTCTGGACA TCGAAACAGC CACTCATGCA
181  GGAAAGCAGA TAGTGGAGCA GATTCTGGAA AAGGAATCAG ATGAGGCACT TAAAATGACC
241  ATTGCCTCTG TTCCTACTTC ACGCTACTTA ACTGACATGA CTCTTGATGA GATGTCAAGA
301  GACTGGTTCA TGCTCATGCC CAAGCAGAAA GTAACAGGCT CCCTATGTAT AAGAATGGAC
361  CAGGCAATCA TGGATAAGAA CATCATACTT AAAGCAAACT TTAGTGTGAT TTTCGAAAGG
421  CTGGAAACAC TAATACTACT TAGAGCCTTC ACCGAAGAAG GAGCAGTCGT TGGCGAAATT
481  TCACCATTAC CTTCTCTTCC AGGACATACT AATGAGGATG TCAAAAATGC AATTGGGGTC
541  CTCATCGGAG GACTTAAATG GAATGATAAT ACGGTTAGAA TCTCTGAAAC TCTACAGAGA
601  TTCGCTTGGA GAAGCAGTCA TGAGAATGGG AGACCTTCAT TCCCTCCAAA GCAGAAACGA
661  AAAATGGAGA GAACAATTGA GCCAGAAATT TGA
```

NS1 amino acid sequence (unassigned): 230 aa

```
1    MDSNTVSSFQ VDCFLWHVRK RFADQELGDA PFLDRLRRDQ KSLRGRGSTL GLDIETATHA
61   GKQIVEQILE KESDEALKMT IASVPTSRYL TDMTLDEMSR DWFMLMPKQK VTGSLCIRMD
121  QAIMDKNIIL KANFSVIFER LETLILLRAF TEEGAVVGEI SPLPSLPGHT NEDVKNAIGV
181  LIGGLKWNDN TVRISETLQR FAWRSSHENG RPSFPPKQKR KMERTIEPEI *
```

APPENDIX 15: Sequences of 13 representative EIVs – Part 9

- **A/equine/Kentucky/1999**

NS segment (MF182443): 890 bp

1	AGCAAAAGCA	GGGTGACAAA	AACATAATGG	ATTCCAACAC	TGTGTCAAGC	TTTCAGGTAG
61	ACTGTTTTCT	TTGGCATGTC	CGCAAACGAT	TCGCAGACCA	AGAACTGGGT	GATGCCCCAT
121	TCCTTGACCG	GCTTCGCCGA	GACCAGAAGT	CCCTAAGGGG	AAGAGGTAGC	ACTCTTGGTC
181	TGGACATCGA	AACAGCCACT	CATGCAGGAA	AGCAGATAGT	GGAGCAGATT	CTGGAAAAGG
241	AATCAGATGA	GGCACTTAAA	ATGACCATTG	CCTCTGTTCC	TACTTCACGC	TACTTAACTG
301	ACATGACTCT	TGATGAGATG	TCAAGAGACT	GGTTCATGCT	CATGCCCAAG	CAAAAAGTAA
361	CAGGCTCCCT	ATGTATAAGA	ATGGACCAGG	CAATCATGGA	TAAGAACATC	ATACTTAAAG
421	CAAACTTTAG	TGTGATTKTC	GAAAGGCTGG	AAACACTAAT	ACTACTTAGA	GCCTTCACCG
481	AAGAAGGAGC	AGTCGTTGGC	GAAATTTCAC	CATTACCTTC	TCTTCCAGGA	CATACTAATG
541	AGGATGTCAA	AAATGCAATT	GGGGTCCTCA	TCGGAGGACT	TAAATGGAAT	GATAATACGG
601	TTAGAATCTC	TGAAACTCTA	CAGAGATTCG	CTTGGAGAAG	CAGTCATGAG	AATGGGAGAC
661	CTTCATTCCC	TTCAAAGCAG	AAATGAAAAA	TGGAGAGAAC	AATTGAGCCA	GAAATTTGAA
721	GAAATAAGAT	GGTTGATTGA	AGAAGTGCGA	CATAGATTGA	AAAATACAGA	GAATAGTTTT
781	GAACAAATAA	CATTTATGCA	AGCCTTACAA	CTATTGCTTG	AAGTAGAACA	AGAGATAAGA
841	ACTTTCTCGT	TTCAGCTTAT	TTAATGATAA	AAAACACCCT	TGTTTCTACT	

NS1 CDS (unassigned): 660 bp

1	ATGGATTCCA	ACACTGTGTC	AAGCTTTCAG	GTAGACTGTT	TTCTTTGGCA	TGTCCGCAAA
61	CGATTCGCAG	ACCAAGAACT	GGGTGATGCC	CCATTCCTTG	ACCGGCTTCG	CCGAGACCAG
121	AAGTCCCTAA	GGGGAAGAGG	TAGCACTCTT	GGTCTGGACA	TCGAAACAGC	CACTCATGCA
181	GGAAAGCAGA	TAGTGGAGCA	GATTCTGGAA	AAGGAATCAG	ATGAGGCACT	TAAAATGACC
241	ATTGCCTCTG	TTCCTACTTC	ACGCTACTTA	ACTGACATGA	CTCTTGATGA	GATGTCAAGA
301	GACTGGTTCA	TGCTCATGCC	CAAGCAAAAA	GTAACAGGCT	CCCTATGTAT	AAGAATGGAC
361	CAGGCAATCA	TGGATAAGAA	CATCATACTT	AAAGCAAACT	TTAGTGTGAT	TTTCGAAAGG
421	CTGGAAACAC	TAATACTACT	TAGAGCCTTC	ACCGAAGAAG	GAGCAGTCGT	TGGCGAAATT
481	TCACCATTAC	CTTCTCTTCC	AGGACATACT	AATGAGGATG	TCAAAAATGC	AATTGGGGTC
541	CTCATCGGAG	GACTTAAATG	GAATGATAAT	ACGGTTAGAA	TCTCTGAAAC	TCTACAGAGA
601	TTCGCTTGGA	GAAGCAGTCA	TGAGAATGGG	AGACCTTCAT	TCCCTTCAAA	GCAGAAATGA

NS1 amino acid sequence (unassigned): 219 aa

1	MDSNTVSSFQ	VDCFLWHVRK	RFADQELGDA	PFLDRLRRDQ	KSLRGRGSTL	GLDIETATHA
61	GKQIVEQILE	KESDEALKMT	IASVPTSRYL	TDMTLDEMSR	DWFMLMPKQK	VTGSLCIRMD
121	QAIMDKNIIL	KANFSVIFER	LETLILLRAF	TEEGAVVGEI	SPLPSLPGHT	NEDVKNAIGV
181	LIGGLKWNDN	TVRISETLQR	FAWRSSHENG	RPSFPSKQK		

APPENDIX 16: Sequences of 13 representative EIVs – Part 10

- **A/equine/Kentucky/5/2002**

NS segment (DQ124185): 890 bp

```
1    AGCAAAAGCA GGGTGACAAA AACATAATGG ATTCCAACAC TGTGTCAAGC TTTCAGGTAG
61   ACTGTTTTCT TTGGCATGTC CGCAAACGAT TCGCAGACCA AGAACTGGGT GATGCCCCAT
121  TCCTTGACCG GCTTCGCCGA GACCAGAAGT CCCTAAGGGG AAGAGGTATC ACTCTTGGTC
181  TGGACATCGA AACAGCCACT CATGCAGGAA AGCAGATAGT GGAGCAGATT CTGGAAAAGG
241  AATCAGATGA GGCACTTAAA ATGACCATTG CCTCTGTTCC TACTTCACGC TACTTAACTG
301  ACATGACTCT TGATGAGATG TCAAGAGACT GGTTCATGCT CATGCCCAAG CAAAAAGTAA
361  CAGGCTCCCT ATGTATAAGA ATGGACCAGG CAATCATGGA TAAGAACATC ATACTTAAAG
421  CAAACTTTAG TGTGATTTTC GAAAGGCTGG AAACACTAAT ACTACTTAGA GCCTTCACCG
481  AAGAAGGAGC AGTCGTTGGC GAAATTTCAC CATTACCTTC TCTTCCAGGA CATACTAATG
541  AGGATGTCAA AAATGCAATT GGGGTCCTCA TCGGAGGACT TAAATGGAAT GATAATACGG
601  TTAGAATCTC TGAAACTCTA CAGAGATTCG CTTGGAGAAG CAGTCATGAG AATGGGAGAC
661  CTTCATTCCC TTCAAAGCAG AAATGAAAAA TGGAGAGAAC AATTAAGCCA GAAATTTGAA
721  GAAATAAGAT GGTTGATTGA AGAAGTGCGA CATAGATTGA AAAATACAGA AAATAGTTTT
781  GAACAAATAA CATTTATGCA AGCCTTACAA CTATTGCTTG AAGTAGAACA AGAGATAAGA
841  ACTTTCTCGT TTCAGCTTAT TTAATGATAA AAAACACCCT TGTTTCTACT
```

NS1 CDS (ABA42429): 660 bp

```
1    ATGGATTCCA ACACTGTGTC AAGCTTTCAG GTAGACTGTT TTCTTTGGCA TGTCCGCAAA
61   CGATTCGCAG ACCAAGAACT GGGTGATGCC CCATTCCTTG ACCGGCTTCG CCGAGACCAG
121  AAGTCCCTAA GGGGAAGAGG TATCACTCTT GGTCTGGACA TCGAAACAGC CACTCATGCA
181  GGAAAGCAGA TAGTGGAGCA GATTCTGGAA AAGGAATCAG ATGAGGCACT TAAAATGACC
241  ATTGCCTCTG TTCCTACTTC ACGCTACTTA ACTGACATGA CTCTTGATGA GATGTCAAGA
301  GACTGGTTCA TGCTCATGCC CAAGCAAAAA GTAACAGGCT CCCTATGTAT AAGAATGGAC
361  CAGGCAATCA TGGATAAGAA CATCATACTT AAAGCAAACT TTAGTGTGAT TTTCGAAAGG
421  CTGGAAACAC TAATACTACT TAGAGCCTTC ACCGAAGAAG GAGCAGTCGT TGGCGAAATT
481  TCACCATTAC CTTCTCTTCC AGGACATACT AATGAGGATG TCAAAAATGC AATTGGGGTC
541  CTCATCGGAG GACTTAAATG GAATGATAAT ACGGTTAGAA TCTCTGAAAC TCTACAGAGA
601  TTCGCTTGGA GAAGCAGTCA TGAGAATGGG AGACCTTCAT TCCCTTCAAA GCAGAAATGA
```

NS1 amino acid sequence (ABA42429): 219 aa

```
1    MDSNTVSSFQ VDCFLWHVRK RFADQELGDA PFLDRLRRDQ KSLRGRGITL GLDIETATHA
61   GKQIVEQILE KESDEALKMT IASVPTSRYL TDMTLDEMSR DWFMLMPKQK VTGSLCIRMD
121  QAIMDKNIIL KANFSVIFER LETLILLRAF TEEGAVVGEI SPLPSLPGHT NEDVKNAIGV
181  LIGGLKWNDN TVRISETLQR FAWRSSHENG RPSFPSKQK
```

APPENDIX 17: Sequences of 13 representative EIVs – Part 11

- **A/equine/Newmarket/5/2003**

NS segment (FJ375209): 890 bp

1	AGCAAAAGCA	GGGTGACAAA	AACATAATGG	ATTCCAACAC	TGTGTCAAGC	TTTCAGGTAG
61	ACTGTTTTCT	TTGGCATGTC	CGCAAACGAT	TCGCAGACCA	AGAACTGGGT	GATGCCCCAT
121	TCCTTGACCG	GCTTCGCCGA	GACCAGAAGT	CCCTAAGGGG	AAGAGGTATC	ACTCTTGGTC
181	TGGACATCGA	AACAGCCACT	CATGCAGGAA	AGCAGATAGT	GGAGCAGATT	CTGGAAAAGG
241	AATCAGATGA	GGCACTTAAA	ATGACCATTG	CCTCTATTCC	TACTTCACGC	TACTTAACTG
301	ACATGACTCT	TGATGAGATG	TCAAGAGACT	GGTTCATGCT	CATGCCCAAG	CAAAAAGTAA
361	CAGGCTCCCT	ATGTATAAGA	ATGGACCAGG	CAATCATGGA	TAAAAACATC	ATACTTAAAG
421	CAAACTTTAG	TGTGATTTTC	GAAAGGCTGG	AAACACTAAT	ACTACTTAGA	GCCTTCACCG
481	AAGAAGGAGC	AGTCGTTGGC	GAAATTTCAC	CATTACCTTC	TCTTCCAGGA	CATACTAATG
541	AGGATGTCAA	AAATGCAATT	GGGGTCCTCA	TCGGAGGACT	TAAATGGAAT	GATAATACGG
601	TTAGAATCTC	TGAAACTCTA	CAGAGATTCG	CTTGGAGAAG	CAGTCATGAG	AATGGGAGAC
661	CTTCATTCCC	TTCAAAGCAG	AAATGAAAAA	TGGAGAGAAC	AATTAAGCCA	GAAATTTGAA
721	GAAATAAGAT	GGTTGATTGA	AGAAGTGCGA	CATAGATTGA	AAAATACAGA	AAATAGTTTT
781	GAACAAATAA	CATTTATGCA	AGCCTTACAA	CTATTGCTTG	AAGTAGAACA	AGAGATAAGA
841	ACTTTCTCGT	TTCAGCTTAT	TTAATGATAA	AAAACACCCT	TGTTTCTACT	

NS1 CDS (ACI48802): 660 bp

1	ATGGATTCCA	ACACTGTGTC	AAGCTTTCAG	GTAGACTGTT	TTCTTTGGCA	TGTCCGCAAA
61	CGATTCGCAG	ACCAAGAACT	GGGTGATGCC	CCATTCCTTG	ACCGGCTTCG	CCGAGACCAG
121	AAGTCCCTAA	GGGGAAGAGG	TATCACTCTT	GGTCTGGACA	TCGAAACAGC	CACTCATGCA
181	GGAAAGCAGA	TAGTGGAGCA	GATTCTGGAA	AAGGAATCAG	ATGAGGCACT	TAAAATGACC
241	ATTGCCTCTA	TTCCTACTTC	ACGCTACTTA	ACTGACATGA	CTCTTGATGA	GATGTCAAGA
301	GACTGGTTCA	TGCTCATGCC	CAAGCAAAAA	GTAACAGGCT	CCCTATGTAT	AAGAATGGAC
361	CAGGCAATCA	TGGATAAAAA	CATCATACTT	AAAGCAAACT	TTAGTGTGAT	TTTCGAAAGG
421	CTGGAAACAC	TAATACTACT	TAGAGCCTTC	ACCGAAGAAG	GAGCAGTCGT	TGGCGAAATT
481	TCACCATTAC	CTTCTCTTCC	AGGACATACT	AATGAGGATG	TCAAAAATGC	AATTGGGGTC
541	CTCATCGGAG	GACTTAAATG	GAATGATAAT	ACGGTTAGAA	TCTCTGAAAC	TCTACAGAGA
601	TTCGCTTGGA	GAAGCAGTCA	TGAGAATGGG	AGACCTTCAT	TCCCTTCAAA	GCAGAAATGA

NS1 amino acid sequence (ACI48802): 219 aa

1	MDSNTVSSFQ	VDCFLWHVRK	RFADQELGDA	PFLDRLRRDQ	KSLRGRGITL	GLDIETATHA
61	GKQIVEQILE	KESDEALKMT	IASIPTSRYL	TDMTLDEMSR	DWFMLMPKQK	VTGSLCIRMD
121	QAIMDKNIIL	KANFSVIFER	LETLILLRAF	TEEGAVVGEI	SPLPSLPGHT	NEDVKNAIGV
181	LIGGLKWNDN	TVRISETLQR	FAWRSSHENG	RPSFPSKQK		

APPENDIX 18: Sequences of 13 representative EIVs – Part 12

- **A/equine/Ohio/1/2003**

NS segment (DQ124186): 890 bp

```
1    AGCAAAAGCA GGGTGACAAA AACATAATGG ATTCCAACAC TGTGTCAAGC TTTCAGGTAG
61   ACTGTTTTCT TTGGCATGTC CGCAAACGAT TCGCAGACCA AGAACTGGGT GATGCCCCAT
121  TCCTTGACCG GCTTCGCCGA GACCAGAAGT CCCTAAGGGG AAGAGGTAGC ACTCTTGGTC
181  TGGACATCGA AACAGCCACT CATGCAGGAA AGCAGATAGT GGAGCAGATT CTGGAAAAGG
241  AATCAGATGA GGCACTTAAA ATGACCATTG CCTCTGTTCC TACTTCACGC TACTTAACTG
301  ACATGACTCT TGATGAGATG TCAAGAGACT GGTTCATGCT CATGCCCAAG CAAAAAGTAA
361  CAGGCTCCCT ATGTATAAGA ATGGACCAGG CAATCATGGA TAAGAACATC ATACTTAAAG
421  CAAACTTTAG TGTGATTTTC GAAAGGCTGG AAACACTAAT ACTACTTAGA GCCTTCACCG
481  AAGAAGGAGC AGTCGTTGGC GAAATTTCAC CATTACCTTC TCTTCCAGGA CATACTAATG
541  AGGATGTCAA AAATGCAATT GGGGTCCTCA TCGGAGGACT TAAATGGAAT GATAATACGG
601  TTAGAATCTC TGAAACTCTA CAGAGATTCG CTTGGAGAAG CAGTCATGAG AATGGGAGAC
661  CTTCATTCCC TTCAAAGCAG AAATGAAAAA TGGAGAGAAC AATTAAGCCA GAAATTTGAA
721  GAAATAAGAT GGTTGATTGA AGAAGTGCGA CATAGATTGA AAAATACAGA AAATAGTTTT
781  GAACAAATAA CATTTATGCA AGCCTTACAA CTATTGCTTG AAGTAGAACA AGAGATAAGA
841  ACTTTCTCGT TTCAGCTTAT TTAATGATAA AAAACACCCT TGTTTCTACT
```

NS1 CDS (ABA42431): 660 bp

```
1    ATGGATTCCA ACACTGTGTC AAGCTTTCAG GTAGACTGTT TTCTTTGGCA TGTCCGCAAA
61   CGATTCGCAG ACCAAGAACT GGGTGATGCC CCATTCCTTG ACCGGCTTCG CCGAGACCAG
121  AAGTCCCTAA GGGGAAGAGG TAGCACTCTT GGTCTGGACA TCGAAACAGC CACTCATGCA
181  GGAAAGCAGA TAGTGGAGCA GATTCTGGAA AAGGAATCAG ATGAGGCACT TAAAATGACC
241  ATTGCCTCTG TTCCTACTTC ACGCTACTTA ACTGACATGA CTCTTGATGA GATGTCAAGA
301  GACTGGTTCA TGCTCATGCC CAAGCAAAAA GTAACAGGCT CCCTATGTAT AAGAATGGAC
361  CAGGCAATCA TGGATAAGAA CATCATACTT AAAGCAAACT TTAGTGTGAT TTTCGAAAGG
421  CTGGAAACAC TAATACTACT TAGAGCCTTC ACCGAAGAAG GAGCAGTCGT TGGCGAAATT
481  TCACCATTAC CTTCTCTTCC AGGACATACT AATGAGGATG TCAAAAATGC AATTGGGGTC
541  CTCATCGGAG GACTTAAATG GAATGATAAT ACGGTTAGAA TCTCTGAAAC TCTACAGAGA
601  TTCGCTTGGA GAAGCAGTCA TGAGAATGGG AGACCTTCAT TCCCTTCAAA GCAGAAATGA
```

NS1 amino acid sequence (ABA42431): 219 aa

```
1    MDSNTVSSFQ VDCFLWHVRK RFADQELGDA PFLDRLRRDQ KSLRGRGSTL GLDIETATHA
61   GKQIVEQILE KESDEALKMT IASVPTSRYL TDMTLDEMSR DWFMLMPKQK VTGSLCIRMD
121  QAIMDKNIIL KANFSVIFER LETLILLRAF TEEGAVVGEI SPLPSLPGHT NEDVKNAIGV
181  LIGGLKWNDN TVRISETLQR FAWRSSHENG RPSFPSKQK
```

APPENDIX 19: Sequences of 13 representative EIVs – Part 13

- **A/equine/Mongolia/3/2013**

NS segment (MF182459): 890 bp

```
1    AGCAAAAGCA GGGTGACAAA AACATAATGG ATTCCAACAC TGTGTCAAGC TTTCAGGTAG
61   ACTGTTTTCT TTGGCATGTC CGTAAACGAT TCGCAGACCA AGAACTGGGT GATGCCCCAT
121  TCCTTGACCG GCTTCGCCGA GACCAGAAGT CCCTAAGAAG AAGAGGTATC ACTCTTGGTC
181  TGGACATCGA AACAGCCACT CATGCAGGAA AGCAGATAGT GGAGCAGATT CTGGAAAAGG
241  AATCAGATGA GGCACTTAAA ATGACCATTG CCTCTATTCC TACTTCACGT TACTTAACTG
301  ACATGACTCT TGATGAGATG TCAAGAGACT GGTTCATGCT CATGCCCAAG CAAAAAGTAA
361  CAGGCTCCCT ATGTATAAGA ATGGACCAGG CAATTATGGA TAAGAACATC ATACTTAAAG
421  CAAACTTTAG TGTGATTTTC GAAAGGCTGG AAACACTAAT ACTACTTAGA GCCTTCACCG
481  AAGAAGGAGC AGTCGTTGGC GAAATTTCAC CATTACCTTC TCTTCCAGGA CATACTAATG
541  AGGATGTCAA AAATGCAATT GGGGTCCTCA TCGGAGGACT TAAATGGAAT GATAATACGG
601  TTAGAATCTC TGAAACTCTA CAGAGATTCG CTTGGAGAAG CAGTCATGAG AATGGGAGAC
661  CTTCATTCCC TTCAAAGCAG AAATGAAAAA TGGAGAGAAC AATTAAGCCA GAAATTTGAA
721  GAAATAAGAT GGTTGATTGA AGAAGTGCGA CATAGATTGA AAAATACAGA AAATAGTTTT
781  GAACAAATAA CATTTATGCA AGCCTTACAA CTATTGCTTG AAGTGGAACA AGAGATAAGA
841  ACTTTCTCGT TTCAGCTTAT TTAATGATAA AAAACACCCT TGTTTCTACT
```

NS1 CDS (unassigned): 660 bp

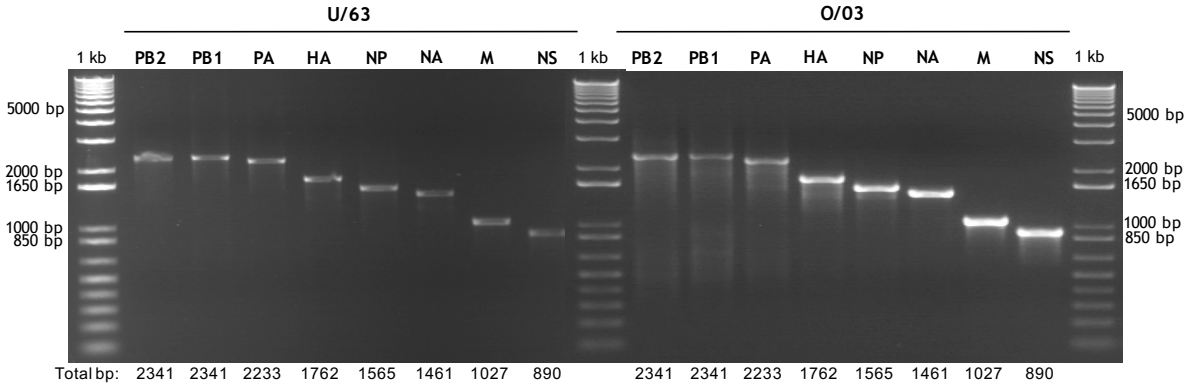
```
1    ATGGATTCCA ACACTGTGTC AAGCTTTCAG GTAGACTGTT TTCTTTGGCA TGTCCGTAAA
61   CGATTCGCAG ACCAAGAACT GGGTGATGCC CCATTCCTTG ACCGGCTTCG CCGAGACCAG
121  AAGTCCCTAA GAAGAAGAGG TATCACTCTT GGTCTGGACA TCGAAACAGC CACTCATGCA
181  GGAAAGCAGA TAGTGGAGCA GATTCTGGAA AAGGAATCAG ATGAGGCACT TAAAATGACC
241  ATTGCCTCTA TTCCTACTTC ACGTTACTTA ACTGACATGA CTCTTGATGA GATGTCAAGA
301  GACTGGTTCA TGCTCATGCC CAAGCAAAAA GTAACAGGCT CCCTATGTAT AAGAATGGAC
361  CAGGCAATTA TGGATAAGAA CATCATACTT AAAGCAAACT TTAGTGTGAT TTTCGAAAGG
421  CTGGAAACAC TAATACTACT TAGAGCCTTC ACCGAAGAAG GAGCAGTCGT TGGCGAAATT
481  TCACCATTAC CTTCTCTTCC AGGACATACT AATGAGGATG TCAAAAATGC AATTGGGGTC
541  CTCATCGGAG GACTTAAATG GAATGATAAT ACGGTTAGAA TCTCTGAAAC TCTACAGAGA
601  TTCGCTTGGA GAAGCAGTCA TGAGAATGGG AGACCTTCAT TCCCTTCAAA GCAGAAATGA
```

NS1 amino acid sequence (unassigned): 219 aa

```
1    MDSNTVSSFQ VDCFLWHVRK RFADQELGDA PFLDRLRRDQ KSLRRRGITL GLDIETATHA
61   GKQIVEQILE KESDEALKMT IASIPTSRYL TDMTLDEMSR DWFMLMPKQK VTGSLCIRMD
121  QAIMDKNIIL KANFSVIFER LETLILLRAF TEEGAVVGEI SPLPSLPGHT NEDVKNAIGV
181  LIGGLKWNDN TVRISETLQR FAWRSSHENG RPSFPSKQK
```

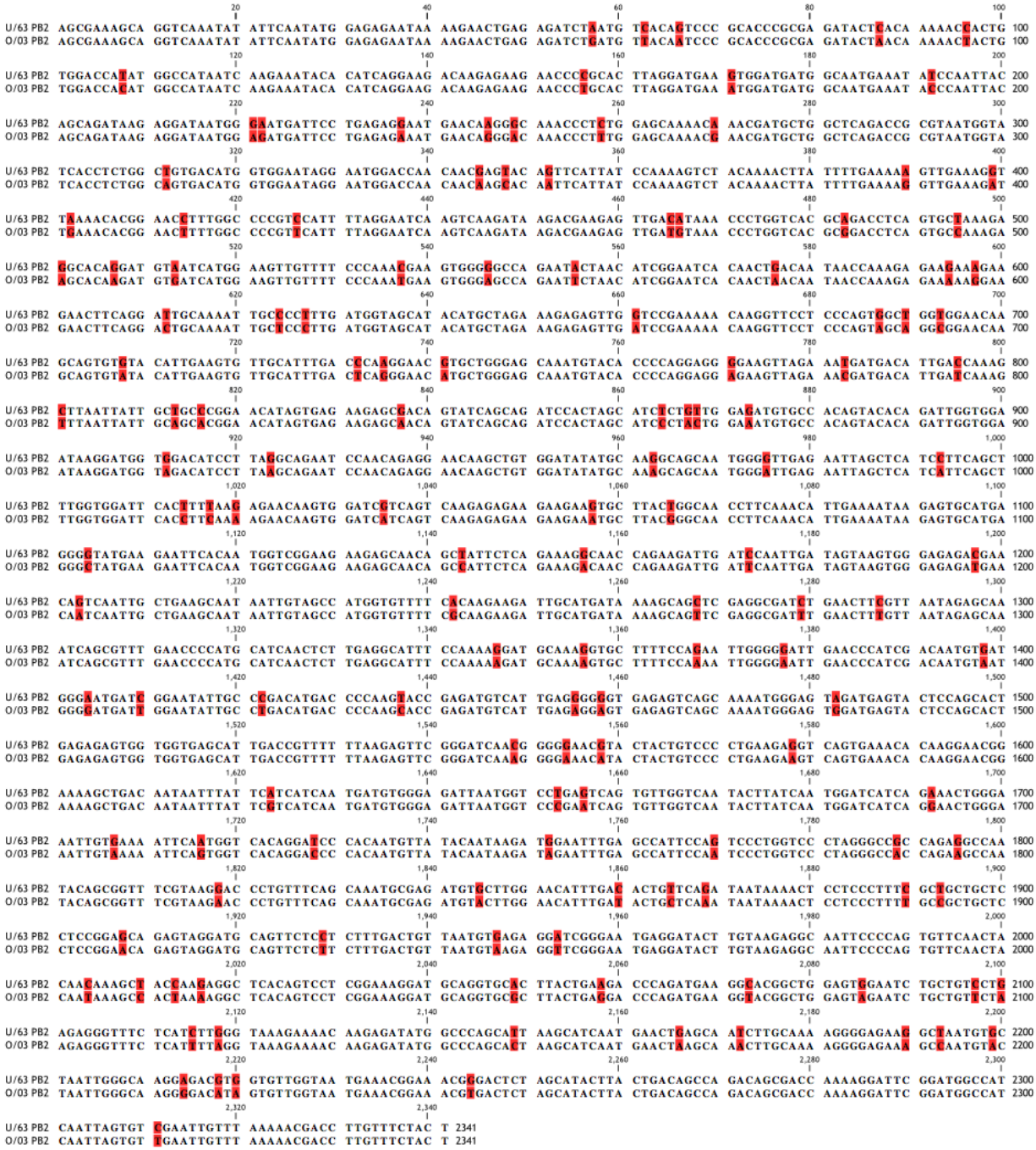
APPENDIX 20: Ambisense plasmid stocks of U/63 and O/03 viruses.

Two sets of 8 ambisense expression plasmids encoding for the viral segments of U/63 and O/03, were grown as described in the method section. A specific PCR was done for each segment and the PCR products were run on a 1% agarose gel to check for their respective size. The number of base pairs (total bp) found in each segment is indicated below the gel, under the corresponding band.



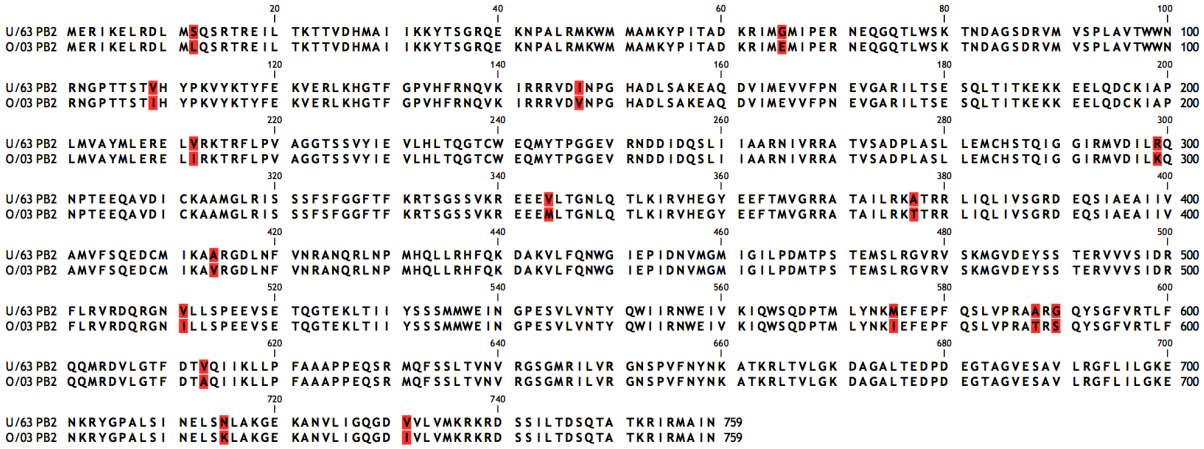
APPENDIX 21: Sequencing of U/63 & O/03 PB2 segments – Part 1

Nucleotide sequence alignment



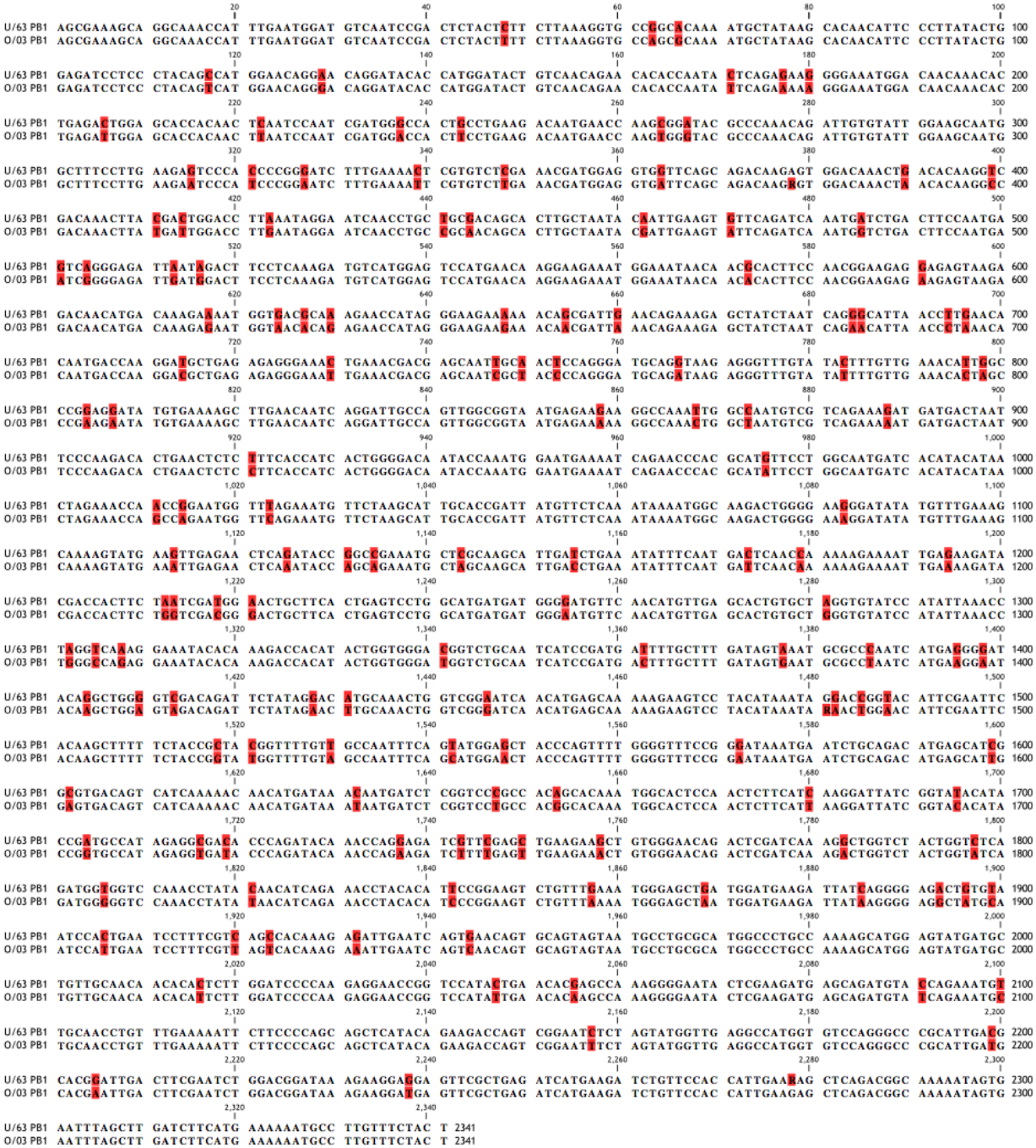
APPENDIX 22: Sequencing of U/63 & O/03 PB2 segments – Part 2

Amino acid sequence alignment



APPENDIX 23: Sequencing of U/63 & O/03 PB1 segments – Part 1

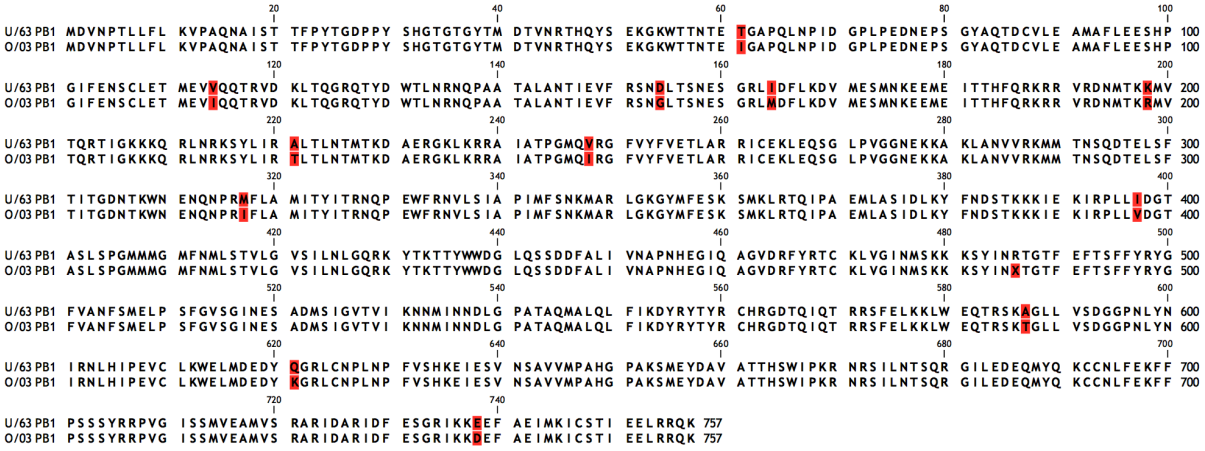
Nucleotide sequence alignment



APPENDIX 24: Sequencing of U/63 & O/03 PB1 segments – Part 2

Amino acid sequence alignment

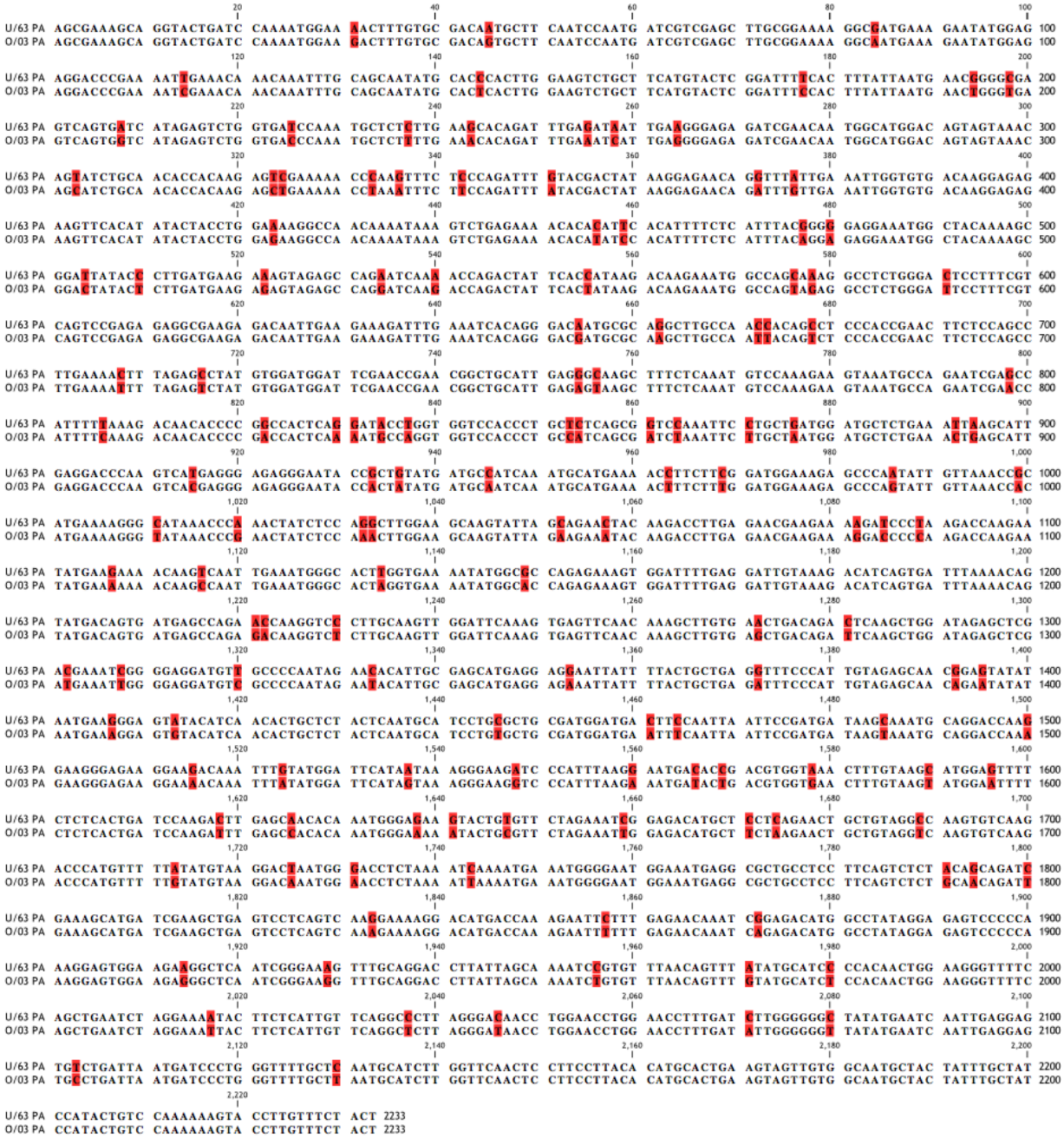
- **PB1**



- **PB1-F2**

APPENDIX 25: Sequencing of U/63 & O/03 PA segments – Part 1

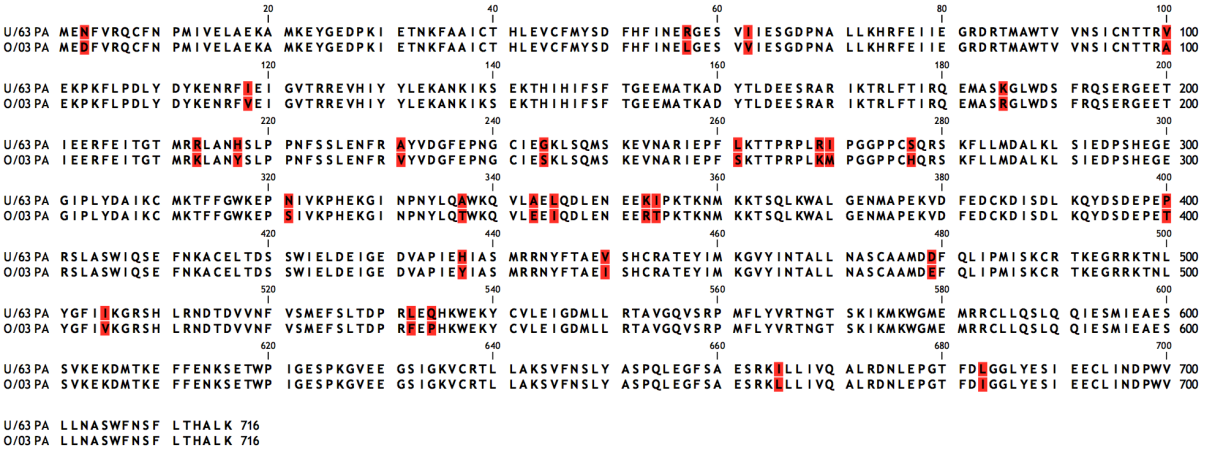
Nucleotide sequence alignment



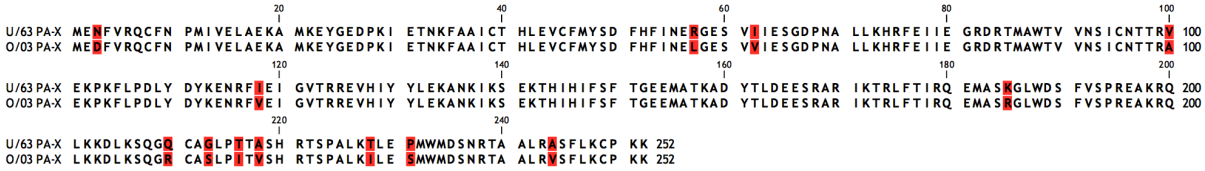
APPENDIX 26: Sequencing of U/63 & O/03 PA segments – Part 2

Amino acid sequence alignment

➢ PA

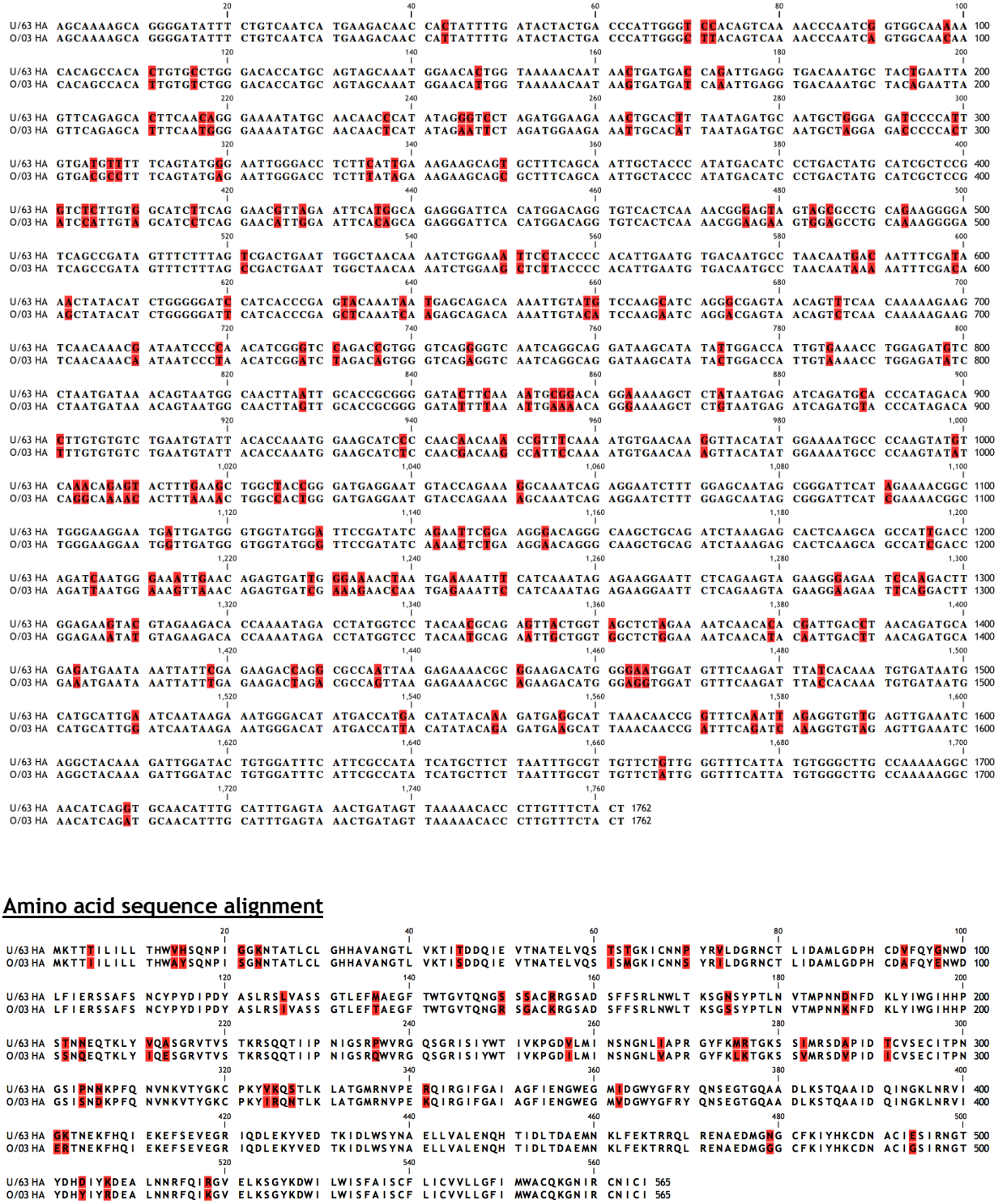


➢ PA-X



APPENDIX 27: Sequencing of U/63 & O/03 HA segments

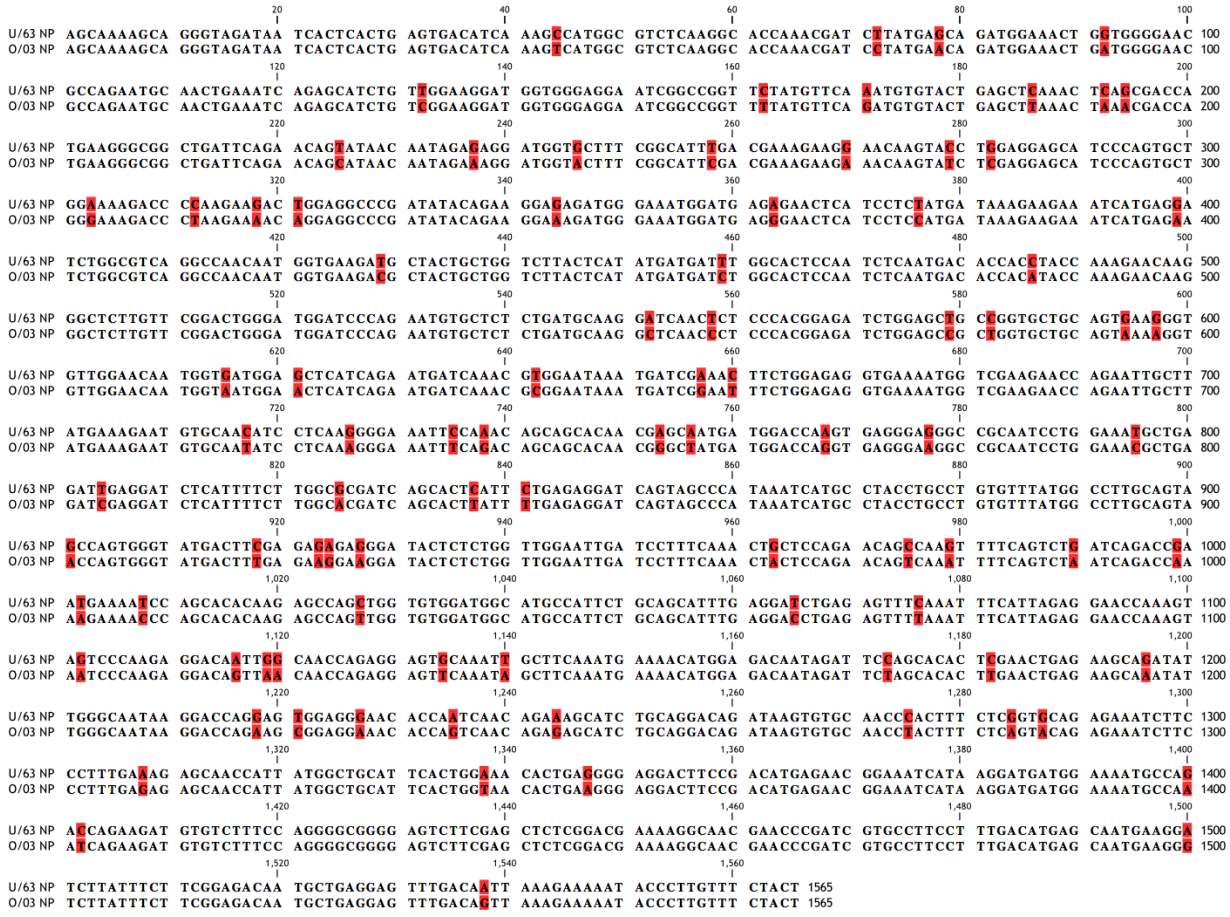
Nucleotide sequence alignment



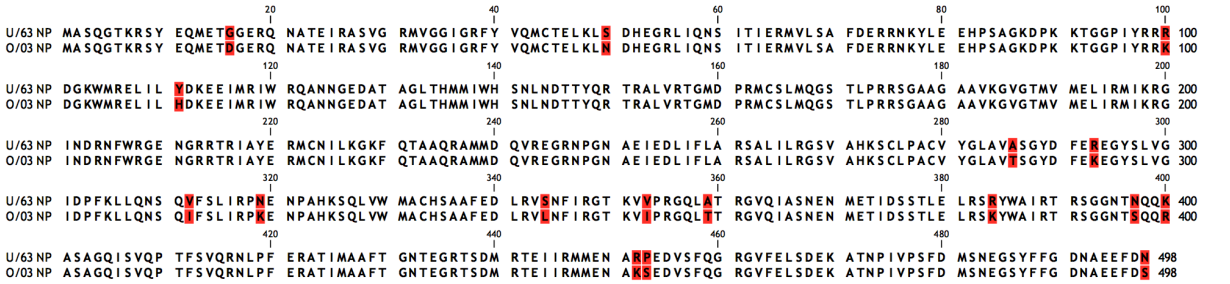
Amino acid sequence alignment

APPENDIX 28: Sequencing of U/63 & O/03 NP segments

Nucleotide sequence alignment

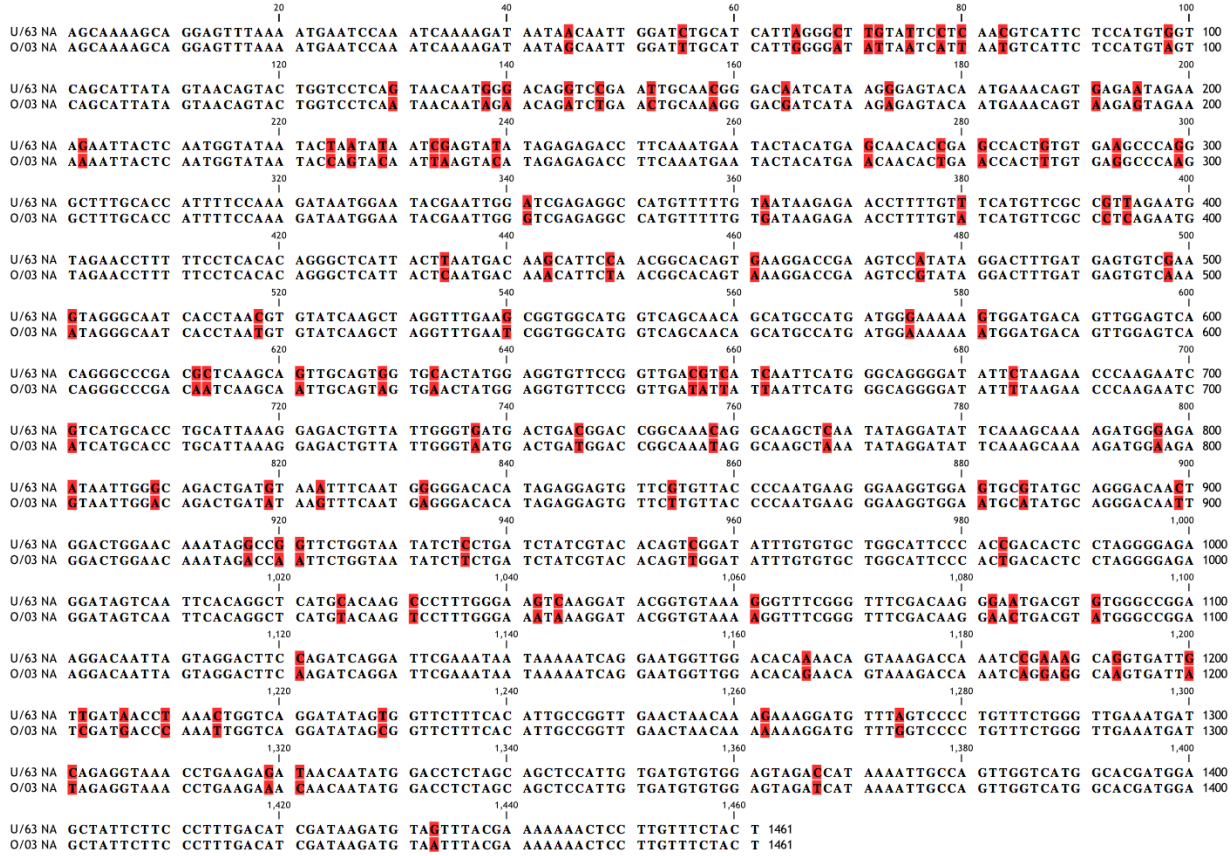


Amino acid sequence alignment

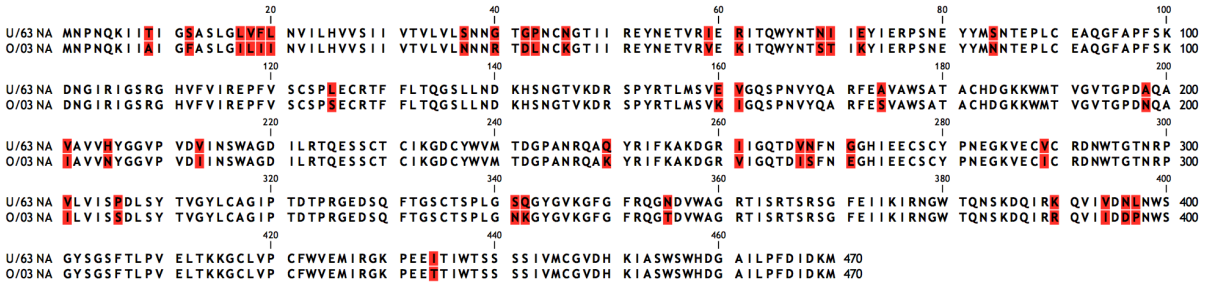


APPENDIX 29: Sequencing of U/63 & O/03 NA segments

Nucleotide sequence alignment

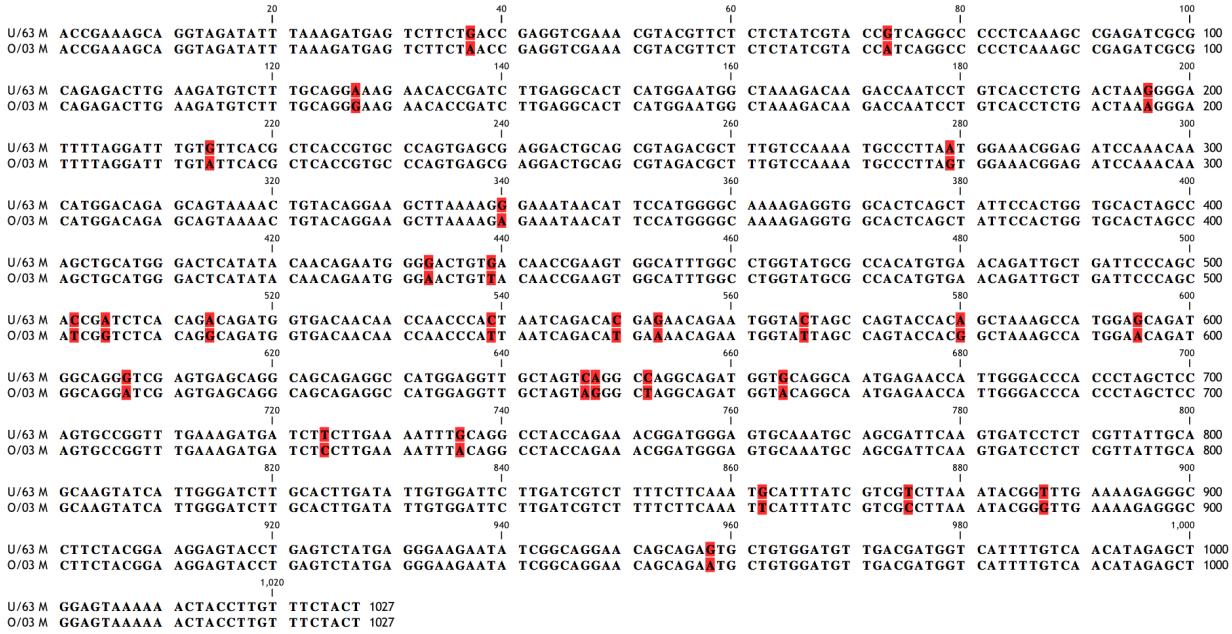


Amino acid sequence alignment



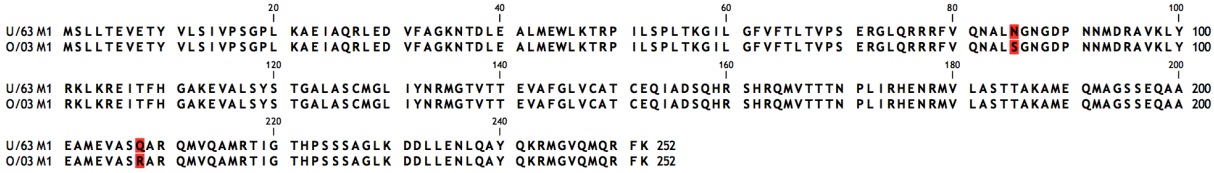
APPENDIX 30: Sequencing of U/63 & O/03 M segments

Nucleotide sequence alignment



Amino acid sequence alignment

- **M1**



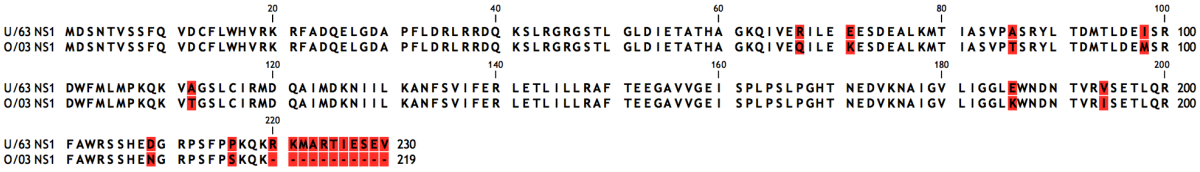
- **M2**

APPENDIX 31: Alignment of U/63 and O/03 NS segment

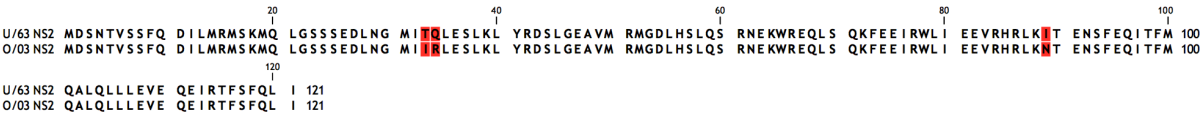
Nucleotide sequence alignment

Amino acid sequence alignment

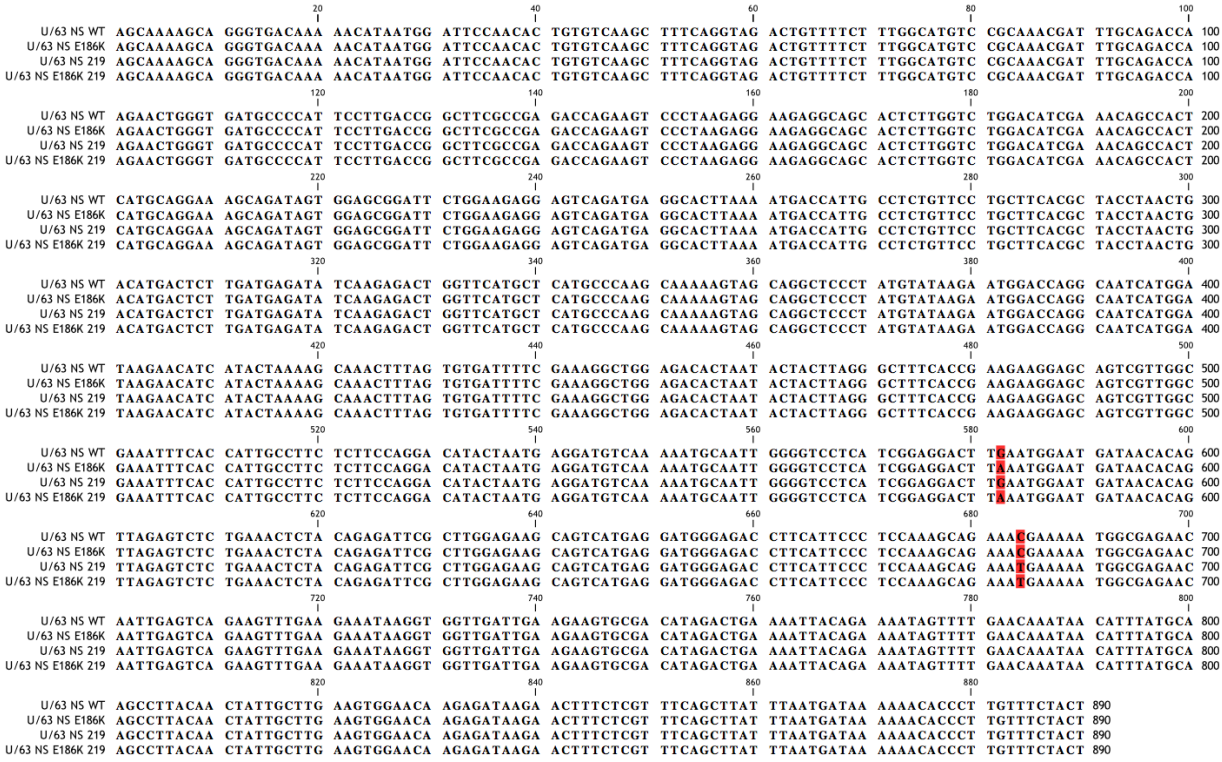
NS1



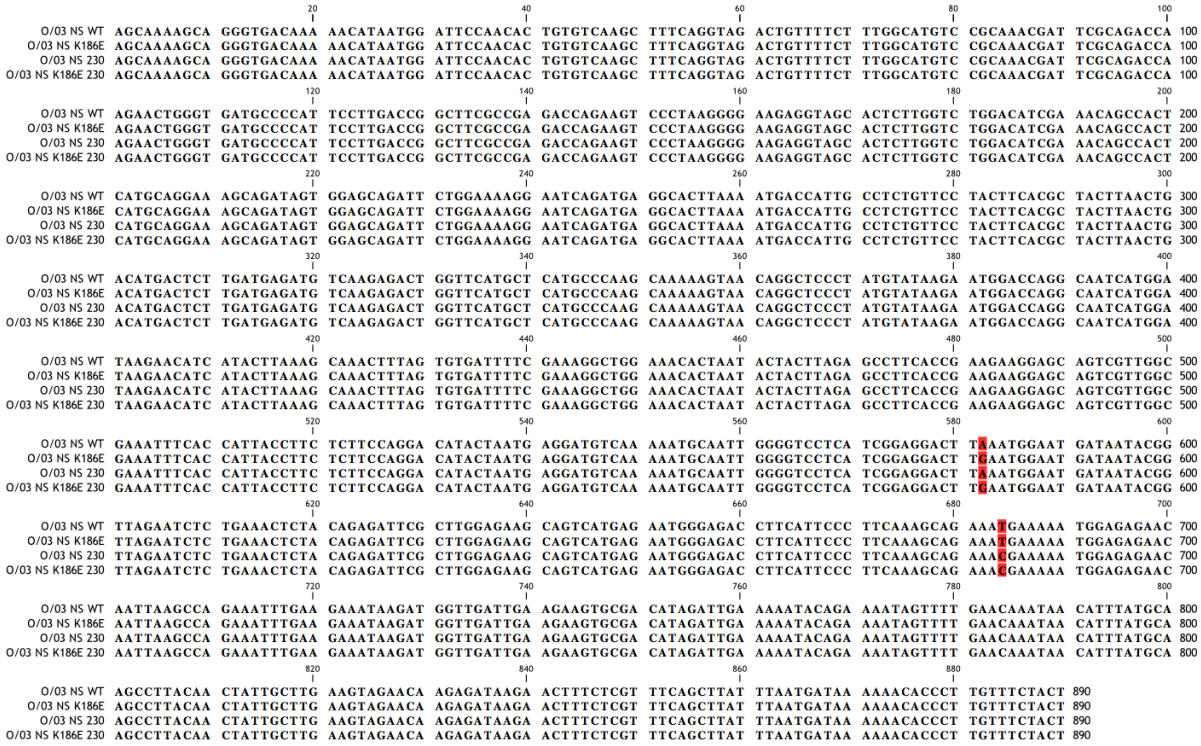
NS2



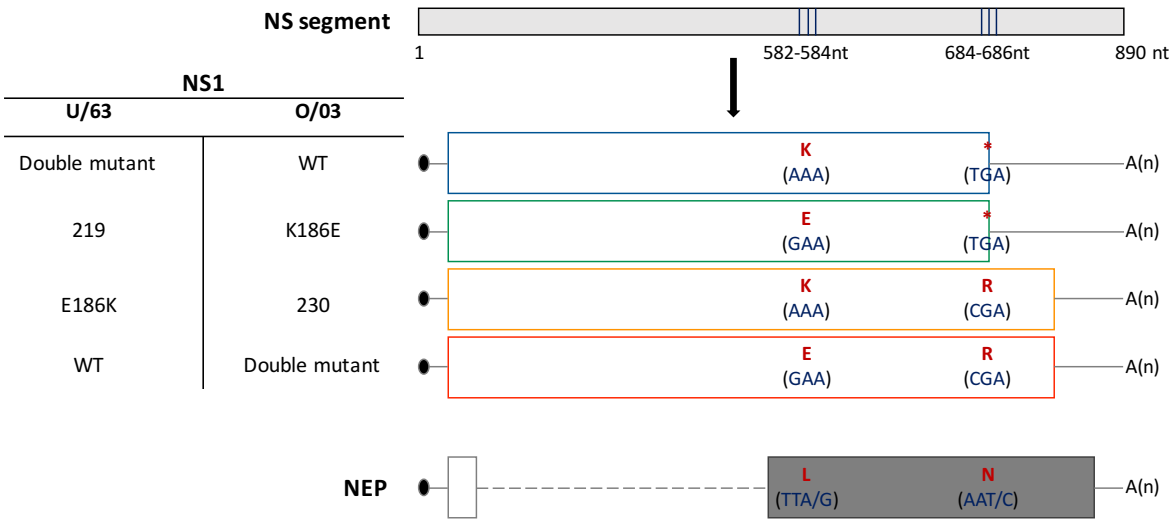
APPENDIX 32: Alignment of U/63 NS WT and mutant inserts in ambisense plasmid



APPENDIX 33: Alignment of O/03 NS WT and mutant inserts in ambisense plasmid



APPENDIX 34: Mutation of codon 186 and 220 in NS ambisense plasmid



APPENDIX 35: List of DEGs INO/03-infected samples – Part 1

		O/03-WT		O/03-K186E		O/03-230		O/03-K186E-230	
Gene ID	**Gene name**	**Rank**	**Log2FC**	**Rank**	**Log2FC**	**Rank**	**Log2FC**	**Rank**	**Log2FC**
ENSECAG00000010153	IFIT3	Up15	2.8169	Up2	5.21902	Up3	3.73116	Up3	4.28147
ENSECAG00000027377	5_8S_rRNA	Up39	1.31487	Up12	2.86602	Up14	1.38835	Up15	1.73267
ENSECAG00000023135	LYRM7	Up53	1.06639	Up57	0.740132	Up43	0.941237	Up34	1.03292
ENSECAG00000004574	ENSECAG00000004574	Up41	1.23114	Up37	1.14136	Up53	0.879453	Up37	0.941438
ENSECAG00000004433	IFIT1	Up104	0.771454	Up7	4.13469	Up5	2.51653	Up9	3.1852
ENSECAG00000008899	LSS	Down137	-0.624569	Down31	-0.820577	Down11	-0.962401	Down5	-1.00929
ENSECAG00000009202	TYSND1	Down27	-0.907853	Down18	-0.942333	Down32	-0.740013	Down48	-0.615778
ENSECAG00000004572	CH25H	Down12	-1.13537	Down10	-1.04437	Down9	-1.01427	Down8	-0.977711
ENSECAG00000025867	U3	Up17	2.40357	Up15	2.44815			Up5	3.37152
ENSECAG00000019095	ENSECAG00000019095	Up46	1.16439	Up49	0.866737			Up56	0.704857
ENSECAG00000027676	ENSECAG00000027676	Up57	1.02918	Up35	1.1478			Up27	1.16402
ENSECAG00000016730	ENSECAG00000016730	Up85	0.84377	Up43	0.966973			Up39	0.893828
ENSECAG00000003015	SERPINB2	Up151	0.676015	Up58	0.684345			Up83	0.61405
ENSECAG00000019165	CENPK	Up177	0.649676	Up65	0.581247			Up45	0.823367
ENSECAG00000016459	ENSECAG00000016459	Up190	0.642261	Up64	0.594166			Up89	0.609109
ENSECAG00000004107	DDIT4	Down175	-0.586586	Down103	-0.608796			Down25	-0.716045
ENSECAG00000000429	POLRMT	Down157	-0.604098	Down82	-0.650021			Down47	-0.61954
ENSECAG00000019125	FASN	Down155	-0.605124	Down32	-0.819688			Down27	-0.713131
ENSECAG00000012635	BRAT1	Down146	-0.611826	Down86	-0.641948			Down29	-0.692893
ENSECAG00000004005	MESDC1	Down134	-0.627071	Down37	-0.785754			Down49	-0.608919
ENSECAG00000018860	ZFP36L1	Down113	-0.653149	Down84	-0.64632			Down51	-0.60197
ENSECAG00000001989	SESN2	Down88	-0.698381	Down94	-0.62295			Down32	-0.682113
ENSECAG00000000607	TRIB3	Down50	-0.800805	Down38	-0.781903			Down11	-0.833385
ENSECAG00000013176	MEGF6	Down39	-0.842357	Down80	-0.665181			Down37	-0.664524
ENSECAG00000009989	CHAC1	Down29	-0.901175	Down27	-0.849656			Down23	-0.72979
ENSECAG00000011234	COL7A1	Down26	-0.926848	Down89	-0.636524			Down33	-0.675143
ENSECAG00000011475	ENSECAG00000011475	Down17	-1.05965	Down9	-1.08052			Down9	-0.924415
ENSECAG00000009190	TMSB4X	Up58	1.02308			Up15	1.3809	Up68	0.658129
ENSECAG00000021345	POLR2K	Up51	1.08059			Up22	1.25657	Up72	0.646725
ENSECAG00000009383	TBCA	Up77	0.919505			Up13	1.39272	Up79	0.62422
ENSECAG00000013756	C8orf59	Up82	0.869592			Up87	0.730823	Up50	0.734264
ENSECAG00000020689	SPA17	Up197	0.633619			Up135	0.642672	Up94	0.600385
ENSECAG00000014624	SPX	Up243	0.586334			Up119	0.671073	Up87	0.61023
ENSECAG00000008688	ENSECAG00000008688	Down100	-0.673453	Down101	-0.611868	Down2	-1.23933		
ENSECAG00000019576	ENSECAG00000019576	Up3	inf					Up1	inf
ENSECAG00000020819	AGER	Up26	1.65782					Up18	1.53619
ENSECAG00000005401	ENSECAG00000005401	Up32	1.48633					Up22	1.29573
ENSECAG00000012348	EFNB2	Up33	1.45556					Up30	1.08431
ENSECAG00000014338	EGR1	Up35	1.38557					Up23	1.27712
ENSECAG00000015342	CXCL8	Up40	1.30399					Up31	1.08159
ENSECAG00000007744	RSRP1	Up44	1.1724					Up36	0.979854
ENSECAG00000017181	PTGS2	Up67	0.954572					Up41	0.868289
ENSECAG00000001334	C3orf14	Up75	0.922307					Up64	0.671459
ENSECAG00000005903	ENSECAG00000005903	Up80	0.899233					Up44	0.825845
ENSECAG00000003755	PIGW	Up87	0.833383					Up92	0.600669
ENSECAG00000007070	TPRKB	Up101	0.80597					Up69	0.657097
ENSECAG00000025162	ATP5EP2	Up115	0.752403					Up78	0.631377
ENSECAG00000024911	ITGB3BP	Up118	0.746261					Up82	0.61876
ENSECAG00000010468	CHCHD7	Up120	0.743429					Up98	0.597031
ENSECAG00000027699	ENSECAG00000027699	Up129	0.714441					Up42	0.82849
ENSECAG00000018495	SSB	Up130	0.71181					Up93	0.60062
ENSECAG00000019053	FAM175A	Up136	0.7075					Up62	0.681148
ENSECAG00000026862	GCSH	Up145	0.69445					Up91	0.606816
ENSECAG00000011843	ENSECAG00000011843	Up153	0.673509					Up66	0.664985
ENSECAG00000018894	RPL22L1	Up161	0.665869					Up67	0.664411
ENSECAG00000027675	ND4	Up170	0.653594					Up52	0.726651
ENSECAG00000011465	ISCA1	Up184	0.645685					Up84	0.613693
ENSECAG00000020965	CMC1	Up191	0.641895					Up99	0.594351
ENSECAG00000027684	ND1	Up200	0.626917					Up49	0.735849
ENSECAG00000020387	MRPS18C	Up204	0.625188					Up59	0.6987
ENSECAG00000016939	RPS24	Up212	0.619026					Up81	0.621074
ENSECAG00000017538	LY96	Up215	0.614693					Up75	0.634798
ENSECAG00000020416	ARPC3	Up225	0.604788					Up96	0.598332
ENSECAG00000020144	UBXN2A	Up237	0.590855					Up100	0.594333
ENSECAG00000023484	SKA1	Up244	0.585847					Up76	0.633546
ENSECAG00000009792	UNC5B	Down160	-0.598728					Down22	-0.734704
ENSECAG00000023637	ADAM33	Down135	-0.626909					Down36	-0.666905
ENSECAG00000007338	SCARF2	Down106	-0.66066					Down54	-0.593362
ENSECAG00000018651	GTPBP3	Down96	-0.682899					Down30	-0.684212
ENSECAG00000003436	APBB3	Down65	-0.736995					Down42	-0.640553
ENSECAG00000007112	ENSECAG00000007112	Down41	-0.831601					Down40	-0.652355
ENSECAG00000012059	OPLAH	Down40	-0.838903					Down35	-0.670492
ENSECAG00000023782	SLC16A13	Down10	-1.20782					Down2	-1.18221

APPENDIX 36: List of DEGs IN O/03-infected samples – Part 2

		O/03-WT		O/03-K186E		O/03-230		O/03-K186E-230	
Gene ID	Gene name	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC
ENSECAG00000019932	ENSECAG00000019932			Up17	2.33603	Up9	1.62176		
ENSECAG00000009543	EPSTI1			Up31	1.21872	Up23	1.25582		
ENSECAG00000014295	PI3			Up40	1.08421	Up44	0.920787		
ENSECAG00000012662	SEMA3A			Down113	-0.596351	Down30	-0.741631		
ENSECAG00000014601	GADD45A			Down111	-0.597874	Down42	-0.678195		
ENSECAG00000010554	MSMO1			Down79	-0.666701	Down3	-1.18586		
ENSECAG00000011271	IGFBP4			Down74	-0.670052	Down43	-0.673379		
ENSECAG00000020860	LSMEM1	Up13	3.19098	Up14	2.51879				
ENSECAG00000022299	LRRC66	Up28	1.59154	Up33	1.20033				
ENSECAG00000004180	HSPA6	Up31	1.50962	Down5	-1.31121				
ENSECAG00000026824	LYRM4	Up96	0.817396	Up53	0.764381				
ENSECAG00000026993	CEBPZOS	Up116	0.749826	Up56	0.740731				
ENSECAG00000013685	TNXB	Down170	-0.588493	Down115	-0.594876				
ENSECAG00000006963	NACC2	Down156	-0.604823	Down97	-0.620019				
ENSECAG00000021411	TM7SF2	Down147	-0.611736	Down47	-0.73979				
ENSECAG00000024890	CCDC120	Down141	-0.616803	Down56	-0.72215				
ENSECAG00000012684	SWSAP1	Down138	-0.623099	Down98	-0.615043				
ENSECAG00000010395	RIN1	Down130	-0.636608	Down99	-0.614772				
ENSECAG00000010506	PLEKHH3	Down124	-0.643436	Down125	-0.58471				
ENSECAG00000018552	CDKN2A	Down120	-0.646724	Down95	-0.621609				
ENSECAG00000001181	CBX4	Down119	-0.647958	Down119	-0.591799				
ENSECAG00000006501	FJX1	Down118	-0.649401	Down121	-0.590583				
ENSECAG00000013443	AKNA	Down117	-0.649643	Down118	-0.592903				
ENSECAG00000016720	OBSCN	Down111	-0.654409	Down120	-0.59137				
ENSECAG00000012169	ATP5D	Down110	-0.656075	Down93	-0.623178				
ENSECAG00000020898	RELB	Down104	-0.663394	Down104	-0.606461				
ENSECAG00000011704	TNFAIP2	Down103	-0.667399	Down29	-0.834622				
ENSECAG00000015306	ATG4D	Down97	-0.682022	Down81	-0.658705				
ENSECAG00000006302	SLC25A29	Down93	-0.688544	Down53	-0.729061				
ENSECAG00000013820	KMT5C	Down86	-0.702321	Down39	-0.779386				
ENSECAG00000021748	ENSECAG00000021748	Down80	-0.705971	Down23	-0.864304				
ENSECAG00000003315	IRS2	Down79	-0.707088	Down85	-0.64472				
ENSECAG00000021641	SNX21	Down75	-0.718118	Down46	-0.744673				
ENSECAG00000016064	SYDE1	Down74	-0.720635	Down54	-0.725743				
ENSECAG00000017274	KRT13	Down72	-0.722815	Down100	-0.613584				
ENSECAG00000019910	PDLIM2	Down69	-0.728577	Down61	-0.710742				
ENSECAG00000023023	OSGIN1	Down57	-0.762054	Down43	-0.765127				
ENSECAG00000013201	MMP17	Down51	-0.794045	Down30	-0.821235				
ENSECAG00000013281	IRGQ	Down47	-0.804184	Down19	-0.928778				
ENSECAG00000014596	GYLTL1B	Down44	-0.815025	Down15	-0.974631				
ENSECAG00000003075	BORCS6	Down43	-0.818891	Down11	-1.03005				
ENSECAG00000019726	ENSECAG00000019726	Down42	-0.819503	Down12	-1.01958				
ENSECAG00000003591	TPGS1	Down38	-0.84294	Down67	-0.694915				
ENSECAG00000020219	ENSECAG00000020219	Down36	-0.852069	Down20	-0.920106				
ENSECAG00000006082	ENSECAG00000006082	Down34	-0.863856	Down22	-0.866239				
ENSECAG00000010241	SPSB4	Down33	-0.867167	Down70	-0.691728				
ENSECAG00000021543	CDKN1A	Down25	-0.933939	Down17	-0.954175				
ENSECAG00000004462	HIC1	Down21	-0.989624	Down63	-0.701717				
ENSECAG00000005956	TPBGL	Down20	-1.00602	Down13	-1.01796				
ENSECAG00000019043	SSC5D	Down15	-1.07302	Down21	-0.888949				
ENSECAG00000000154	SDHAF1	Down14	-1.07843	Down3	-1.38192				
ENSECAG00000011185	BTBD19	Down8	-1.22584	Down14	-1.00361				
ENSECAG00000006098	ANKRD9	Down2	-1.95687	Down1	-1.94064				
ENSECAG00000015180	NEXN					Up21	1.26305	Up58	0.702545
ENSECAG00000016232	IL22RA1					Up91	0.728298	Up48	0.774814
ENSECAG00000008852	ABCA1					Up160	0.603089	Up53	0.717703
ENSECAG00000010720	ELOVL6					Down28	-0.752294	Down17	-0.747894
ENSECAG00000014658	SCD					Down17	-0.898684	Down55	-0.590867
ENSECAG00000005106	FIBIN					Down15	-0.917468	Down18	-0.740236
ENSECAG00000008637	HMGCS1					Down14	-0.935547	Down24	-0.72822
ENSECAG00000015140	LIPG					Down12	-0.942586	Down7	-0.989717
ENSECAG00000010664	GNRH1	Up2	inf			Up1	inf		
ENSECAG00000025973	U1	Up8	inf			Up2	inf		
ENSECAG00000016983	ANKRD1	Up47	1.14911			Up29	1.1145		
ENSECAG00000011446	ENSECAG00000011446	Up72	0.932911			Up19	1.26822		
ENSECAG00000025120	SBSN	Up86	0.843737			Up34	1.06734		
ENSECAG00000018861	ENSECAG00000018861	Up91	0.827969			Up100	0.713307		
ENSECAG00000020428	ENSECAG00000020428	Up110	0.760672			Up37	0.989556		
ENSECAG00000013843	TUBE1	Up114	0.752447			Up62	0.828848		
ENSECAG00000021551	UTP11L	Up119	0.74427			Up88	0.730403		
ENSECAG00000018544	SPDL1	Up121	0.740986			Up164	0.599705		
ENSECAG00000000651	SPINK5	Up133	0.709617			Up47	0.910982		
ENSECAG00000000708	SARNP	Up134	0.709311			Up150	0.627748		

APPENDIX 37: List of DEGs IN O/03-infected samples – Part 3

		O/03-WT		O/03-K186E		O/03-230		O/03-K186E-230	
Gene ID	Gene name	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC
ENSECAG00000002607	WBP5	Up144	0.695514			Up27	1.14543		
ENSECAG00000010777	FAM107B	Up154	0.672125			Up38	0.989344		
ENSECAG00000007912	HMGN4	Up156	0.670885			Up56	0.869776		
ENSECAG00000015007	SKA2	Up162	0.665233			Up90	0.729377		
ENSECAG00000020945	MNS1	Up193	0.639925			Up33	1.07119		
ENSECAG00000019024	FAM204A	Up196	0.636851			Up157	0.604718		
ENSECAG00000003783	ENSECAG00000003783	Up199	0.630491			Up138	0.639737		
ENSECAG00000026937	SNRNP48	Up202	0.626273			Up147	0.629964		
ENSECAG00000006773	CIR1	Up221	0.609636			Up45	0.914679		
ENSECAG00000024519	SPC25	Up232	0.598761			Up179	0.585074		
ENSECAG00000026920	HSBP1	Up236	0.59354			Up111	0.6827		
ENSECAG00000011948	FIGF	Down81	-0.704733			Down19	-0.841851		
ENSECAG00000024260	TMEM160	Down52	-0.793587			Down13	-0.937601		
ENSECAG00000001022	UCN2	Down13	-1.08918			Up61	0.835742		
ENSECAG00000024167	DDX60			Up3	4.63181			Up6	3.31012
ENSECAG00000007881	IFIH1			Up6	4.40685			Up4	3.5855
ENSECAG00000010185	ENSECAG00000010185			Up16	2.33874			Up14	1.90592
ENSECAG00000025328	ENSECAG00000025328			Up18	2.28563			Up13	1.90771
ENSECAG00000000968	ENSECAG00000000968			Up21	2.10439			Up33	1.0504
ENSECAG00000027056	ENSECAG00000027056			Up22	1.95999			Up19	1.51838
ENSECAG00000001481	SAMD9L			Up27	1.48532			Up40	0.876975
ENSECAG00000021220	ENSECAG00000021220			Up29	1.34581			Up46	0.789085
ENSECAG00000012331	PARP9			Up38	1.12343			Up60	0.692215
ENSECAG00000011726	EIF2AK2			Up41	1.04597			Up47	0.781161
ENSECAG00000001399	SAMD9			Up51	0.816776			Up95	0.600358
ENSECAG00000010055	RSAD2			Up52	0.780623			Up51	0.730225
ENSECAG00000002264	ISL1			Down117	-0.593754			Down38	-0.664083
ENSECAG00000010566	ENSECAG00000010566			Down65	-0.699738			Down50	-0.608559
ENSECAG00000019557	LDLR			Down58	-0.712427			Down26	-0.715306
ENSECAG00000011206	MVD			Down40	-0.779099			Down20	-0.736368
ENSECAG00000011439	DHX58			Up4	4.61608	Up4	3.09973	Up8	3.19554
ENSECAG00000009020	IFI44			Up5	4.44821	Up6	2.17917	Up7	3.27331
ENSECAG00000009618	IFI44L			Up8	3.82991	Up7	1.77632	Up10	2.80855
ENSECAG00000001324	ISG15			Up9	3.01778	Up8	1.74358	Up11	2.04705
ENSECAG00000000284	ENSECAG00000000284			Up10	2.99697	Up11	1.43109	Up16	1.72675
ENSECAG00000011776	MX1			Up11	2.92668	Up12	1.40675	Up12	2.01323
ENSECAG00000012132	ENSECAG00000012132			Up19	2.21372	Up20	1.26429	Up24	1.23645
ENSECAG00000014422	OAS2			Up20	2.11659	Up50	0.897975	Up25	1.21193
ENSECAG00000024429	IRF9			Up23	1.95478	Up10	1.533	Up21	1.36893
ENSECAG00000001598	IFI35			Up25	1.56816	Up70	0.806505	Up29	1.12405
ENSECAG00000014226	ENSECAG00000014226			Up28	1.40464	Up30	1.10198	Up28	1.13741
ENSECAG00000025265	7SK			Up30	1.23285	Up16	1.32949	Up17	1.63899
ENSECAG00000021133	FDFT1			Down108	-0.603939	Down29	-0.745322	Down31	-0.68223
ENSECAG00000023067	MVK			Down57	-0.717767	Down26	-0.753929	Down28	-0.697242
ENSECAG00000016389	ACSS2			Down52	-0.729119	Down23	-0.791528	Down43	-0.636744
ENSECAG00000000714	FAM26D	Up1	inf						
ENSECAG00000022734	SGPP2	Up4	inf						
ENSECAG00000026877	C17orf98	Up5	inf						
ENSECAG00000009955	TEX43	Up6	inf						
ENSECAG00000027680	ENSECAG00000027680	Up7	inf						
ENSECAG00000007414	SV2A	Up9	4.7725						
ENSECAG00000000207	ACTA1	Up10	4.27728						
ENSECAG00000009535	FCER1G	Up11	3.62833						
ENSECAG00000024000	PYGM	Up12	3.57668						
ENSECAG00000009793	TAGLN3	Up14	3.0704						
ENSECAG00000004408	GPR1	Up16	2.79532						
ENSECAG00000008107	FERMT3	Up18	2.34706						
ENSECAG00000009256	CPE	Up19	2.33724						
ENSECAG00000019367	ENSECAG00000019367	Up20	2.03985						
ENSECAG00000016226	ANKRD35	Up21	1.93647						
ENSECAG00000012826	CDO1	Up22	1.93113						
ENSECAG00000006922	ENSECAG00000006922	Up23	1.86167						
ENSECAG00000015232	ZNF619	Up24	1.82611						
ENSECAG00000008044	PPP1R27	Up25	1.76342						
ENSECAG00000018052	ZNF514	Up27	1.59839						
ENSECAG00000015717	TSSK3	Up29	1.54859						
ENSECAG00000018029	THSD1	Up30	1.51379						
ENSECAG00000012898	PNLDC1	Up34	1.39412						
ENSECAG00000007194	CHRNE	Up36	1.3793						
ENSECAG00000009142	ENSECAG00000009142	Up37	1.37497						
ENSECAG00000020764	ENSECAG00000020764	Up38	1.34451						
ENSECAG00000005115	ZNF454	Up42	1.2287						
ENSECAG00000010656	HBEGF	Up43	1.22256						

APPENDIX 38: List of DEGs IN O/03-infected samples – Part 4

		O/03-WT		O/03-K186E		O/03-230		O/03-K186E-230	
Gene ID	Gene name	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC
ENSECAG00000012746	STAR	Up45	1.16959						
ENSECAG00000020952	ENSECAG00000020952	Up48	1.13517						
ENSECAG00000008668	CSF3	Up49	1.11637						
ENSECAG00000012179	RGS2	Up50	1.1118						
ENSECAG00000024298	PTPN22	Up52	1.06985						
ENSECAG00000021582	ENSECAG00000021582	Up54	1.06426						
ENSECAG00000016648	IL5	Up55	1.05992						
ENSECAG00000008723	CASP12	Up56	1.04135						
ENSECAG00000013528	AVIL	Up59	1.01567						
ENSECAG00000018966	FBXO48	Up60	1.01195						
ENSECAG00000014572	ENSECAG00000014572	Up61	0.989668						
ENSECAG00000003681	ENSECAG00000003681	Up62	0.980483						
ENSECAG00000020923	LEAP2	Up63	0.980251						
ENSECAG00000016085	SYCP2	Up64	0.967566						
ENSECAG00000012653	DYRK4	Up65	0.965215						
ENSECAG00000012680	N4BP2L1	Up66	0.961023						
ENSECAG00000008533	ITGA10	Up68	0.951696						
ENSECAG00000024279	FAM200B	Up69	0.945443						
ENSECAG00000013527	ENSECAG00000013527	Up70	0.939975						
ENSECAG00000006567	SLITRK4	Up71	0.939637						
ENSECAG00000010606	CCDC169	Up73	0.930152						
ENSECAG00000023507	GPR19	Up74	0.925208						
ENSECAG00000013748	KRT25	Up76	0.921815						
ENSECAG00000003669	HLA-DMA	Up78	0.909888						
ENSECAG00000026852	ZCCHC10	Up79	0.905548						
ENSECAG00000005249	IQUB	Up81	0.883674						
ENSECAG00000012801	KBTBD3	Up83	0.868281						
ENSECAG00000010700	CREG2	Up84	0.844446						
ENSECAG00000021939	FAM122C	Up88	0.832568						
ENSECAG00000018428	HSPH1	Up89	0.831423						
ENSECAG00000002021	ENSECAG00000002021	Up90	0.829023						
ENSECAG00000011618	EDN1	Up92	0.824924						
ENSECAG00000001667	DEPDC7	Up93	0.823128						
ENSECAG00000002452	CCDC121	Up94	0.820488						
ENSECAG00000013653	TPMT	Up95	0.81773						
ENSECAG00000004698	PTGR1	Up97	0.816365						
ENSECAG00000011003	PDE4B	Up98	0.814737						
ENSECAG00000022568	TMEM154	Up99	0.809303						
ENSECAG00000019742	CCDC126	Up100	0.807428						
ENSECAG00000024185	GNG11	Up102	0.778049						
ENSECAG00000002759	MTRF1L	Up103	0.776082						
ENSECAG00000020892	TLR10	Up105	0.770745						
ENSECAG00000013766	E2F5	Up106	0.766488						
ENSECAG00000016965	SUV39H2	Up107	0.764009						
ENSECAG00000020223	HACD1	Up108	0.763083						
ENSECAG00000020380	COMMD8	Up109	0.761872						
ENSECAG00000004590	TIPIN	Up111	0.758445						
ENSECAG00000001407	CHSY3	Up112	0.756332						
ENSECAG00000013255	ENSECAG00000013255	Up113	0.755364						
ENSECAG00000006869	MAP9	Up117	0.748459						
ENSECAG00000013661	FAM184B	Up122	0.740297						
ENSECAG00000017348	BCAP29	Up123	0.736949						
ENSECAG00000011592	HNMT	Up124	0.734835						
ENSECAG00000012354	LMLN	Up125	0.725545						
ENSECAG00000015510	FGF7	Up126	0.723648						
ENSECAG00000014615	STX2	Up127	0.719371						
ENSECAG00000010953	TNIP3	Up128	0.717209						
ENSECAG00000019097	TANK	Up131	0.710888						
ENSECAG00000016351	ENSECAG00000016351	Up132	0.71022						
ENSECAG00000022765	ZC3H8	Up135	0.709287						
ENSECAG00000022541	ZNF25	Up137	0.706972						
ENSECAG00000007540	CYP7B1	Up138	0.706454						
ENSECAG00000010696	ENSECAG00000010696	Up139	0.705146						
ENSECAG00000016586	NUCB2	Up140	0.704499						
ENSECAG00000010433	CHMP2B	Up141	0.700278						
ENSECAG00000010663	SKIL	Up142	0.698517						
ENSECAG00000016844	MEIOC	Up143	0.696118						
ENSECAG00000009108	C1orf189	Up146	0.693171						
ENSECAG00000006803	SYTL2	Up147	0.69218						
ENSECAG00000024373	ZBTB11	Up148	0.686358						
ENSECAG00000016487	ATP6V1E1	Up149	0.68053						
ENSECAG00000010450	CCDC59	Up150	0.677447						
ENSECAG00000014597	MCTS1	Up152	0.675524						

APPENDIX 39: List of DEGs IN O/03-infected samples – Part 5

		O/03-WT		O/03-K186E		O/03-230		O/03-K186E-230	
Gene ID	Gene name	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC
ENSECAG00000019989	SNRPG	Up155	0.671256						
ENSECAG00000011295	THAP2	Up157	0.668933						
ENSECAG00000017572	ZNF280C	Up158	0.667444						
ENSECAG00000024614	FAM161A	Up159	0.66698						
ENSECAG00000003510	TXNDC9	Up160	0.666237						
ENSECAG00000024815	RFESD	Up163	0.663531						
ENSECAG00000008493	SERPINB10	Up164	0.662478						
ENSECAG00000004919	CACYBP	Up165	0.659535						
ENSECAG00000026897	MRPS36	Up166	0.655053						
ENSECAG00000008428	CEP83	Up167	0.654762						
ENSECAG00000023743	CCDC58	Up168	0.654648						
ENSECAG00000002943	ENSECAG00000002943	Up169	0.654099						
ENSECAG00000024870	CWH43	Up171	0.653246						
ENSECAG00000024704	CETN3	Up172	0.653123						
ENSECAG00000000880	SPATA33	Up173	0.652458						
ENSECAG00000008390	FOXN2	Up174	0.651412						
ENSECAG00000010323	ENSECAG00000010323	Up175	0.651347						
ENSECAG00000016530	LYPLA1	Up176	0.64999						
ENSECAG00000009675	N4BP2	Up178	0.649494						
ENSECAG00000009898	UEVLD	Up179	0.649393						
ENSECAG00000015185	AK6	Up180	0.649205						
ENSECAG00000000546	CCDC62	Up181	0.648597						
ENSECAG00000009776	CEP57L1	Up182	0.646206						
ENSECAG00000024399	FGF12	Up183	0.645883						
ENSECAG00000021026	BCAS2	Up185	0.645444						
ENSECAG00000002513	ERCC6L	Up186	0.642977						
ENSECAG00000013569	LAMTOR3	Up187	0.642613						
ENSECAG00000015459	ENSECAG00000015459	Up188	0.642552						
ENSECAG00000015237	PIBF1	Up189	0.642502						
ENSECAG00000015355	CLK1	Up192	0.640763						
ENSECAG00000018554	ENSECAG00000018554	Up194	0.639487						
ENSECAG00000011862	C11orf1	Up195	0.638822						
ENSECAG00000015376	GLRX2	Up198	0.631247						
ENSECAG00000015464	RPL36AL	Up201	0.626847						
ENSECAG00000005764	ENSECAG00000005764	Up203	0.625927						
ENSECAG00000005178	CCSAP	Up205	0.624865						
ENSECAG00000023869	CRLS1	Up206	0.624073						
ENSECAG00000015765	THAP6	Up207	0.621924						
ENSECAG00000004316	HRSP12	Up208	0.621403						
ENSECAG00000019506	ARRDC3	Up209	0.620906						
ENSECAG00000019146	NDUFAF6	Up210	0.620436						
ENSECAG00000006587	RWDD1	Up211	0.620345						
ENSECAG00000008716	EIF3J	Up213	0.61719						
ENSECAG00000015757	RPL26	Up214	0.616644						
ENSECAG00000018478	NIFK	Up216	0.614659						
ENSECAG00000011393	GKAP1	Up217	0.614196						
ENSECAG00000018693	ENSECAG00000018693	Up218	0.612525						
ENSECAG00000011218	HAT1	Up219	0.610503						
ENSECAG00000017063	TIMM9	Up220	0.610056						
ENSECAG00000014668	ZNF614	Up222	0.608307						
ENSECAG00000009780	UQCRB	Up223	0.607478						
ENSECAG00000012316	DTWD1	Up224	0.605808						
ENSECAG00000007803	LSM3	Up226	0.602689						
ENSECAG00000018635	ZNF674	Up227	0.601725						
ENSECAG00000010661	RPL21	Up228	0.601547						
ENSECAG00000017850	TAF1D	Up229	0.601284						
ENSECAG00000016758	ENSECAG00000016758	Up230	0.601139						
ENSECAG00000019035	LIPT1	Up231	0.60087						
ENSECAG00000008524	SSBIP1	Up233	0.598617						
ENSECAG00000015468	MGAT4A	Up234	0.598338						
ENSECAG00000015373	ENSECAG00000015373	Up235	0.593591						
ENSECAG00000020566	DYNC2LI1	Up238	0.589788						
ENSECAG00000011795	OSGEPL1	Up239	0.589122						
ENSECAG00000024976	OCIAD1	Up240	0.588753						
ENSECAG00000017059	FAM105A	Up241	0.587727						
ENSECAG00000014972	PHF11	Up242	0.587703						
ENSECAG00000009760	GMNN	Up245	0.585393						
ENSECAG00000012399	SUB1	Up246	0.583617						
ENSECAG00000014530	Sep-10	Up247	0.581098						
ENSECAG00000013324	TXN	Up248	0.58067						
ENSECAG00000023217	KLHL17	Down181	-0.580187						
ENSECAG00000011637	SLC1A5	Down180	-0.580647						
ENSECAG00000017955	SLC12A9	Down179	-0.580741						

APPENDIX 40: List of DEGs IN O/03-infected samples – Part 6

		O/03-WT		O/03-K186E		O/03-230		O/03-K186E-230	
Gene ID	Gene name	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC
ENSECAG00000000402	NDUFS7	Down178	-0.580877						
ENSECAG00000022744	RHBDD3	Down177	-0.582825						
ENSECAG00000014263	ENSECAG00000014263	Down176	-0.585909						
ENSECAG00000002668	RHOB	Down174	-0.586719						
ENSECAG00000011565	GMPPA	Down173	-0.587266						
ENSECAG00000008220	ENSECAG00000008220	Down172	-0.587499						
ENSECAG00000012351	FUK	Down171	-0.588028						
ENSECAG00000009148	LPAR2	Down169	-0.589867						
ENSECAG00000022309	DGKQ	Down168	-0.590556						
ENSECAG00000023189	DDAH2	Down167	-0.593593						
ENSECAG00000003398	MARCKS	Down166	-0.593864						
ENSECAG00000024775	PRKD2	Down165	-0.594245						
ENSECAG00000008659	SHISA5	Down164	-0.594512						
ENSECAG00000016677	MFSD3	Down163	-0.597456						
ENSECAG00000019788	WDR83	Down162	-0.598005						
ENSECAG00000026863	EFS	Down161	-0.59856						
ENSECAG00000012961	HSPG2	Down159	-0.600732						
ENSECAG00000014491	NLRC5	Down158	-0.602386						
ENSECAG00000008298	WNT10B	Down154	-0.605397						
ENSECAG00000023401	ZNF414	Down153	-0.60616						
ENSECAG00000021488	DYRK1B	Down152	-0.606423						
ENSECAG00000018474	DDR1	Down151	-0.607061						
ENSECAG00000021964	ASB9	Down150	-0.607786						
ENSECAG00000012006	BCL9L	Down149	-0.608293						
ENSECAG00000023274	LAMA5	Down148	-0.60959						
ENSECAG00000021765	TSPO	Down145	-0.614029						
ENSECAG00000017916	ENSECAG00000017916	Down144	-0.614305						
ENSECAG00000017838	IGFLR1	Down143	-0.615358						
ENSECAG00000015936	ENSECAG00000015936	Down142	-0.615914						
ENSECAG00000012571	ENSECAG00000012571	Down140	-0.618863						
ENSECAG00000020484	TBCC	Down139	-0.621186						
ENSECAG00000021290	MBD6	Down136	-0.625814						
ENSECAG00000002876	FAAP100	Down133	-0.629622						
ENSECAG00000018245	EPHX3	Down132	-0.629753						
ENSECAG00000017322	METTL12	Down131	-0.63042						
ENSECAG00000012500	AP5Z1	Down129	-0.637977						
ENSECAG00000017084	DPYSL4	Down128	-0.638735						
ENSECAG00000016798	MIEF2	Down127	-0.64028						
ENSECAG00000008910	KLRG2	Down126	-0.640342						
ENSECAG00000011354	DHX34	Down125	-0.642514						
ENSECAG00000026952	ENSECAG00000026952	Down123	-0.64348						
ENSECAG00000009314	TMEM132A	Down122	-0.643783						
ENSECAG00000010992	NFATC4	Down121	-0.644947						
ENSECAG00000008387	TSC22D4	Down116	-0.650186						
ENSECAG00000009361	COL5A1	Down115	-0.650422						
ENSECAG00000005577	GPR39	Down114	-0.651523						
ENSECAG00000015448	TAPBP	Down112	-0.653639						
ENSECAG00000001072	TYMP	Down109	-0.657227						
ENSECAG00000021314	COL4A2	Down108	-0.659806						
ENSECAG00000004662	SPTSSA	Down107	-0.660094						
ENSECAG00000017905	LTBP4	Down105	-0.661668						
ENSECAG00000002142	TRIP6	Down102	-0.66844						
ENSECAG00000022526	CDH15	Down101	-0.669839						
ENSECAG00000016695	SLC27A5	Down99	-0.678787						
ENSECAG00000020085	RELL2	Down98	-0.681381						
ENSECAG00000007615	PTRH1	Down95	-0.683101						
ENSECAG00000012759	PLPPR2	Down94	-0.686553						
ENSECAG00000020780	PELP1	Down92	-0.69176						
ENSECAG00000018161	C1QTNF5	Down91	-0.692114						
ENSECAG00000024412	ARHGAP39	Down90	-0.696513						
ENSECAG00000021280	SHISA4	Down89	-0.697033						
ENSECAG00000015814	LRRC73	Down87	-0.698787						
ENSECAG00000015240	PROM2	Down85	-0.702809						
ENSECAG00000006999	PLEKHG4	Down84	-0.702886						
ENSECAG00000009891	ARMC5	Down83	-0.703847						
ENSECAG00000018111	NAPRT	Down82	-0.704411						
ENSECAG00000021034	YDJC	Down78	-0.708416						
ENSECAG00000019381	MEGF8	Down77	-0.713517						
ENSECAG00000013608	KAZALD1	Down76	-0.714613						
ENSECAG00000013419	MNT	Down73	-0.721419						
ENSECAG00000014566	WISP2	Down71	-0.723337						
ENSECAG00000023946	SDC3	Down70	-0.72685						
ENSECAG00000011353	MFSD7	Down68	-0.730068						

APPENDIX 41: List of DEGs IN O/03-infected samples – Part 7

		O/03-WT		O/03-K186E		O/03-230		O/03-K186E-230	
Gene ID	Gene name	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC
ENSECAG00000006116	PPP1R35	Down67	-0.73076						
ENSECAG00000014470	EEPD1	Down66	-0.731131						
ENSECAG00000004551	FOXG1	Down64	-0.741996						
ENSECAG00000011160	TP53INP2	Down63	-0.747131						
ENSECAG00000017501	TOR4A	Down62	-0.747398						
ENSECAG00000006619	ZNF688	Down61	-0.751039						
ENSECAG00000013524	ZNF219	Down60	-0.751192						
ENSECAG00000011907	C19orf54	Down59	-0.751823						
ENSECAG00000002911	KCTD11	Down58	-0.75277						
ENSECAG00000017451	FAM110A	Down56	-0.770603						
ENSECAG00000015612	FAM212A	Down55	-0.785398						
ENSECAG00000007474	SPHK2	Down54	-0.791489						
ENSECAG00000002061	NEURL2	Down53	-0.792894						
ENSECAG00000006058	FUT1	Down49	-0.803043						
ENSECAG00000000440	TMEM8B	Down48	-0.803244						
ENSECAG00000024202	XKR8	Down46	-0.804649						
ENSECAG00000018426	ERF	Down45	-0.812761						
ENSECAG00000023399	GADD45G	Down37	-0.848082						
ENSECAG00000022258	SDSL	Down35	-0.85835						
ENSECAG00000006591	ATG9B	Down32	-0.873						
ENSECAG00000002570	HS3ST1	Down31	-0.884388						
ENSECAG00000007647	COL5A3	Down30	-0.89889						
ENSECAG00000006133	NXPH4	Down28	-0.90606						
ENSECAG00000013867	ENSECAG00000013867	Down24	-0.96001						
ENSECAG00000006208	MBLAC1	Down23	-0.976957						
ENSECAG00000012116	CARNS1	Down22	-0.97915						
ENSECAG00000015692	SERTAD4	Down19	-1.02147						
ENSECAG00000024458	KCTD7	Down18	-1.04533						
ENSECAG00000011161	ENSECAG00000011161	Down16	-1.06533						
ENSECAG00000026812	FANCF	Down11	-1.18548						
ENSECAG00000007684	SLC2A4	Down9	-1.22492						
ENSECAG00000009912	PCSK1N	Down7	-1.32892						
ENSECAG00000018591	CTU1	Down6	-1.35574						
ENSECAG00000025020	FHL1	Down5	-1.35746						
ENSECAG00000009027	WBSCR27	Down4	-1.4755						
ENSECAG00000014282	NPPB	Down3	-1.81488						
ENSECAG00000003656	NANOS1	Down1	-inf						
ENSECAG00000015034	PLAC8			Up1	inf				
ENSECAG00000019309	GBP2			Up13	2.77326				
ENSECAG00000004306	ENSECAG00000004306			Up24	1.70432				
ENSECAG00000008809	OAS3			Up26	1.56128				
ENSECAG00000010036	SLC15A3			Up32	1.21219				
ENSECAG00000016217	RNF213			Up34	1.18669				
ENSECAG00000009686	DTX3L			Up36	1.1458				
ENSECAG00000019411	HERC6			Up39	1.0866				
ENSECAG00000002176	ENSECAG00000002176			Up42	0.996487				
ENSECAG00000017394	ZC3HAV1			Up44	0.95748				
ENSECAG00000020562	AXDND1			Up45	0.929863				
ENSECAG00000013435	OAS1			Up46	0.908969				
ENSECAG00000021989	DDX58			Up47	0.896421				
ENSECAG00000015395	HERC5			Up48	0.889434				
ENSECAG00000012773	CMPK2			Up50	0.862742				
ENSECAG00000009937	UBE2N			Up54	0.754702				
ENSECAG00000001514	ENSECAG00000001514			Up55	0.750024				
ENSECAG00000022601	PARP12			Up59	0.679275				
ENSECAG00000020384	ENSECAG00000020384			Up60	0.638286				
ENSECAG00000006930	ENSECAG00000006930			Up61	0.637825				
ENSECAG00000023195	ENSECAG00000023195			Up62	0.631552				
ENSECAG00000004349	IFIT5			Up63	0.621803				
ENSECAG00000001560	SCO2			Down128	-0.580853				
ENSECAG00000016361	COL18A1			Down127	-0.58119				
ENSECAG00000013728	MFSD2A			Down126	-0.583615				
ENSECAG00000010237	FBXL6			Down124	-0.586447				
ENSECAG00000001342	SCAND1			Down123	-0.588235				
ENSECAG00000005017	FBXW9			Down122	-0.590475				
ENSECAG00000022018	TRIM47			Down116	-0.593891				
ENSECAG00000017795	ENSECAG00000017795			Down114	-0.595276				
ENSECAG00000010553	NGF			Down112	-0.596973				
ENSECAG00000018835	FAM46B			Down110	-0.601717				
ENSECAG00000008333	MSC			Down109	-0.601983				
ENSECAG00000003816	JUNB			Down107	-0.6043				
ENSECAG00000002992	MID1IP1			Down106	-0.604792				
ENSECAG00000001265	TMEM223			Down105	-0.60591				

APPENDIX 42: List of DEGs IN O/03-infected samples – Part 8

		O/03-WT		O/03-K186E		O/03-230		O/03-K186E-230	
Gene ID	Gene name	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC
ENSECAG00000004050	ENSECAG00000004050			Down102	-0.610559				
ENSECAG00000000138	ACVR2A			Down96	-0.620891				
ENSECAG00000024319	ENSECAG00000024319			Down92	-0.634334				
ENSECAG00000019789	STOML1			Down91	-0.635372				
ENSECAG00000006931	ATF4			Down90	-0.636222				
ENSECAG00000004791	TICAM1			Down88	-0.638507				
ENSECAG00000011779	ARID3A			Down87	-0.639476				
ENSECAG00000013794	MRPS12			Down83	-0.649639				
ENSECAG00000014783	ENSECAG00000014783			Down78	-0.668662				
ENSECAG00000021308	APBA3			Down77	-0.669118				
ENSECAG00000016405	REPIN1			Down76	-0.669615				
ENSECAG00000011011	SNAPC2			Down75	-0.669702				
ENSECAG00000024720	ENTHD2			Down73	-0.68113				
ENSECAG00000016486	ARRDC2			Down72	-0.682995				
ENSECAG00000017157	IER3			Down71	-0.685588				
ENSECAG00000020178	WNT9A			Down69	-0.691772				
ENSECAG00000012599	TRMT61A			Down68	-0.694494				
ENSECAG00000024986	ENSECAG00000024986			Down66	-0.697917				
ENSECAG00000004492	RRS1			Down64	-0.701231				
ENSECAG00000000529	MAP1S			Down62	-0.709554				
ENSECAG00000022294	ENSECAG00000022294			Down60	-0.711121				
ENSECAG00000016441	DOHH			Down59	-0.711948				
ENSECAG00000008194	MIDN			Down55	-0.724055				
ENSECAG00000014794	MIIP			Down51	-0.729532				
ENSECAG00000015758	IRX2			Down50	-0.730415				
ENSECAG00000013124	BCL3			Down49	-0.730821				
ENSECAG00000016942	NUDT14			Down48	-0.731937				
ENSECAG00000013787	HMX3			Down45	-0.747514				
ENSECAG00000021909	C14orf80			Down44	-0.751221				
ENSECAG00000019359	DUS3L			Down42	-0.768069				
ENSECAG00000010803	ENSECAG00000010803			Down41	-0.778312				
ENSECAG00000026965	SOCS1			Down36	-0.789464				
ENSECAG00000021887	TRIM62			Down35	-0.803212				
ENSECAG00000022195	ENSECAG00000022195			Down34	-0.804666				
ENSECAG00000004504	C19orf52			Down33	-0.810494				
ENSECAG00000022697	TRABD2B			Down28	-0.843641				
ENSECAG00000013721	SH2D5			Down26	-0.854379				
ENSECAG00000024586	ENSECAG00000024586			Down25	-0.85534				
ENSECAG00000015715	ADAMTS15			Down24	-0.861468				
ENSECAG00000014511	IL11			Down16	-0.96058				
ENSECAG00000012197	SPATA2L			Down8	-1.10335				
ENSECAG00000018561	CDKN2B			Down7	-1.22934				
ENSECAG00000022402	GFOD1			Down6	-1.23897				
ENSECAG00000014921	GPRC5A			Down4	-1.33756				
ENSECAG00000003025	HLX			Down2	-1.63124				
ENSECAG00000003585	RPL19					Up17	1.32346		
ENSECAG00000024426	PDAP1					Up18	1.29205		
ENSECAG00000002024	ENSECAG00000002024					Up24	1.24544		
ENSECAG00000006844	CALD1					Up25	1.19409		
ENSECAG00000006493	PTMA					Up26	1.15249		
ENSECAG00000019300	CARD9					Up28	1.12444		
ENSECAG00000008785	ENSECAG00000008785					Up31	1.09935		
ENSECAG00000016025	SNRNP27					Up32	1.07142		
ENSECAG00000008257	RBMX2					Up35	1.00488		
ENSECAG00000011155	ATP6V1G1					Up36	0.995159		
ENSECAG00000005090	BASP1					Up39	0.962889		
ENSECAG00000009472	CLSPN					Up40	0.955786		
ENSECAG00000014983	HIRIP3					Up41	0.951502		
ENSECAG00000008992	HYPK					Up42	0.944431		
ENSECAG00000021583	LMOD1					Up46	0.911345		
ENSECAG00000008506	HMMR					Up48	0.903499		
ENSECAG00000010638	SREK1IP1					Up49	0.900818		
ENSECAG00000014412	ZMAT2					Up51	0.897111		
ENSECAG00000009621	INCENP					Up52	0.887348		
ENSECAG00000009886	TRAF3IP1					Up54	0.874669		
ENSECAG00000010132	HMGB2					Up55	0.874481		
ENSECAG00000022843	ZC3H13					Up57	0.851282		
ENSECAG00000012666	PSIP1					Up58	0.842963		
ENSECAG00000019845	SPAG7					Up59	0.838775		
ENSECAG00000022972	TCHP					Up60	0.838251		
ENSECAG00000008385	ENSECAG00000008385					Up63	0.828334		
ENSECAG00000015871	CDCA8					Up64	0.821569		
ENSECAG00000008649	CCDC12					Up65	0.821442		

APPENDIX 43: List of DEGs IN O/03-infected samples – Part 9

		O/03-WT		O/03-K186E		O/03-230		O/03-K186E-230	
Gene ID	Gene name	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC
ENSECAG00000000734	HMGA1					Up66	0.812707		
ENSECAG00000012708	DNAJC8					Up67	0.809498		
ENSECAG00000012364	MYEOV2					Up68	0.808633		
ENSECAG00000026810	FAM133B					Up69	0.806944		
ENSECAG00000009987	PCNP					Up71	0.797147		
ENSECAG00000017165	ENSA					Up72	0.797126		
ENSECAG00000012760	LUC7L3					Up73	0.792121		
ENSECAG00000010635	GPS2					Up74	0.784306		
ENSECAG00000021311	NUDC					Up75	0.777538		
ENSECAG00000026980	ENSECAG00000026980					Up76	0.774947		
ENSECAG00000008928	ENSECAG00000008928					Up77	0.769984		
ENSECAG00000024127	CENPF					Up78	0.767092		
ENSECAG00000003089	ENSECAG00000003089					Up79	0.764132		
ENSECAG00000023366	SUDS3					Up80	0.761729		
ENSECAG00000014880	CFAP45					Up81	0.752188		
ENSECAG00000003072	C8orf4					Up82	0.746632		
ENSECAG00000023594	FAM50A					Up83	0.737634		
ENSECAG00000026981	SRRM1					Up84	0.735645		
ENSECAG00000000593	PTN					Up85	0.733931		
ENSECAG00000013485	MFAP1					Up86	0.731433		
ENSECAG00000016693	GOLIM4					Up89	0.729457		
ENSECAG00000003766	GADD45GIP1					Up92	0.727814		
ENSECAG00000013395	RBM25					Up93	0.72709		
ENSECAG00000014949	TPM1					Up94	0.726056		
ENSECAG00000003616	ENSECAG00000003616					Up95	0.725223		
ENSECAG00000017015	MYH11					Up96	0.723533		
ENSECAG00000024809	C1orf131					Up97	0.718484		
ENSECAG00000023288	RNPS1					Up98	0.717031		
ENSECAG00000010941	CCDC150					Up99	0.713529		
ENSECAG00000006666	MURC					Up101	0.712238		
ENSECAG00000026989	PDCD7					Up102	0.711159		
ENSECAG00000008098	TCF7					Up103	0.707057		
ENSECAG00000019569	ENSECAG00000019569					Up104	0.707021		
ENSECAG00000014847	NUPR1					Up105	0.706403		
ENSECAG00000012794	LYAR					Up106	0.705285		
ENSECAG00000014812	KRT23					Up107	0.697669		
ENSECAG00000013815	ENSECAG00000013815					Up108	0.694679		
ENSECAG00000007123	ENSECAG00000007123					Up109	0.68695		
ENSECAG00000019621	ENSECAG00000019621					Up110	0.686873		
ENSECAG00000016947	UPF3B					Up112	0.680315		
ENSECAG00000014069	GOLM1					Up113	0.679488		
ENSECAG00000022323	PSMC3IP					Up114	0.679355		
ENSECAG00000021350	SRSF4					Up115	0.676075		
ENSECAG00000017963	E2F7					Up116	0.675926		
ENSECAG00000021271	ENSECAG00000021271					Up117	0.67362		
ENSECAG00000021276	BRD3					Up118	0.672743		
ENSECAG00000018958	C9orf78					Up120	0.669381		
ENSECAG00000010862	TNIP1					Up121	0.668438		
ENSECAG00000005865	HAUS8					Up122	0.666204		
ENSECAG00000011826	ATPIF1					Up123	0.666041		
ENSECAG00000010367	RAD51AP1					Up124	0.665138		
ENSECAG00000012231	DEK					Up125	0.663819		
ENSECAG00000021097	ACIN1					Up126	0.660863		
ENSECAG00000011870	PYM1					Up127	0.656765		
ENSECAG00000013512	ENSECAG00000013512					Up128	0.655647		
ENSECAG00000014953	KIF15					Up129	0.653919		
ENSECAG00000011337	ODF2					Up130	0.652646		
ENSECAG00000017141	UPF3A					Up131	0.646773		
ENSECAG00000014242	MAD1L1					Up132	0.646476		
ENSECAG00000016495	NUSAP1					Up133	0.646448		
ENSECAG00000026853	CEP152					Up134	0.644139		
ENSECAG00000018476	TPX2					Up136	0.642373		
ENSECAG00000022315	BIN3					Up137	0.641456		
ENSECAG00000002115	ENSECAG00000002115					Up139	0.639271		
ENSECAG00000008466	ENSECAG00000008466					Up140	0.638539		
ENSECAG00000026905	RPL37					Up141	0.637897		
ENSECAG00000018538	PPIG					Up142	0.637641		
ENSECAG00000003357	ING1					Up143	0.635161		
ENSECAG00000012624	POLD3					Up144	0.634971		
ENSECAG00000024902	ENSECAG00000024902					Up145	0.634489		
ENSECAG00000013696	SLU7					Up146	0.633604		
ENSECAG00000001897	SUMO2					Up148	0.628834		
ENSECAG00000015060	PRR11					Up149	0.627963		

APPENDIX 44: List of DEGs IN O/03-infected samples – Part 10

		O/03-WT		O/03-K186E		O/03-230		O/03-K186E-230	
Gene ID	Gene name	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC
ENSECAG00000020610	NDUFB10					Up151	0.61959		
ENSECAG00000015496	ENSECAG00000015496					Up152	0.616784		
ENSECAG00000007487	ENSECAG00000007487					Up153	0.61607		
ENSECAG00000021538	CKAP2L					Up154	0.613215		
ENSECAG00000017105	SEC62					Up155	0.610318		
ENSECAG00000004534	CCDC136					Up156	0.605553		
ENSECAG00000017076	NSRP1					Up158	0.603928		
ENSECAG00000012779	IK					Up159	0.603291		
ENSECAG00000000633	ZWINT					Up161	0.600546		
ENSECAG00000000644	MAP7D3					Up161	0.600546		
ENSECAG00000013640	CEP85					Up162	0.60029		
ENSECAG00000022449	MEAF6					Up163	0.599962		
ENSECAG00000012637	LENG1					Up165	0.598714		
ENSECAG00000018394	MESDC2					Up166	0.597826		
ENSECAG00000022285	ENSECAG00000022285					Up167	0.597506		
ENSECAG00000018438	CEP55					Up168	0.596321		
ENSECAG00000006129	NUP62					Up169	0.595882		
ENSECAG00000014320	PPL					Up170	0.595836		
ENSECAG00000008538	SPECC1					Up171	0.593915		
ENSECAG00000019321	TXLNA					Up172	0.590412		
ENSECAG00000012023	ENSECAG00000012023					Up173	0.590344		
ENSECAG00000017006	CHMP4A					Up174	0.588888		
ENSECAG00000014009	TPM2					Up175	0.587612		
ENSECAG00000005815	LLPH					Up176	0.58692		
ENSECAG00000011537	C14orf119					Up177	0.586732		
ENSECAG00000011778	C11orf58					Up178	0.585521		
ENSECAG00000015878	AMOT					Down62	-0.583045		
ENSECAG00000008577	PPP3CB					Down61	-0.58889		
ENSECAG00000022342	PKIA					Down60	-0.590756		
ENSECAG00000024974	RNF11					Down59	-0.592537		
ENSECAG00000017012	PTEN					Down58	-0.601453		
ENSECAG00000010628	PPP1R14B					Down57	-0.602554		
ENSECAG00000016862	SLC25A33					Down56	-0.608907		
ENSECAG00000023284	ENSECAG00000023284					Down55	-0.619739		
ENSECAG00000003002	CD9					Down54	-0.624481		
ENSECAG00000014232	NR4A2					Down53	-0.625517		
ENSECAG00000011223	HGF					Down52	-0.6264		
ENSECAG00000010571	AGL					Down51	-0.636853		
ENSECAG00000014113	CCNG2					Down50	-0.64024		
ENSECAG00000021307	TGDS					Down49	-0.650154		
ENSECAG00000001318	UBL3					Down48	-0.652025		
ENSECAG00000002481	ENSECAG00000002481					Down47	-0.655159		
ENSECAG00000012577	HOXA10					Down46	-0.658534		
ENSECAG00000008215	TSC22D3					Down45	-0.666727		
ENSECAG00000010827	ENSECAG00000010827					Down44	-0.667793		
ENSECAG00000015783	PTBP2					Down41	-0.686615		
ENSECAG00000005757	ENSECAG00000005757					Down40	-0.690087		
ENSECAG00000019834	CYP39A1					Down39	-0.695967		
ENSECAG00000021330	HMGCR					Down38	-0.700786		
ENSECAG00000012062	EFNA5					Down37	-0.708252		
ENSECAG00000005163	SLC38A6					Down36	-0.712477		
ENSECAG00000011841	FAM174A					Down35	-0.717274		
ENSECAG00000009515	ATRNL1					Down34	-0.729997		
ENSECAG00000016183	STARD4					Down33	-0.739402		
ENSECAG00000008425	LURAP1L					Down31	-0.740235		
ENSECAG00000017787	FLVCR1					Down27	-0.752844		
ENSECAG00000020317	RAP2C					Down25	-0.76238		
ENSECAG00000024992	HSD17B7					Down24	-0.772415		
ENSECAG00000006455	CYP51A1					Down22	-0.803576		
ENSECAG00000000046	FAM13C					Down21	-0.819509		
ENSECAG00000008926	RNF139					Down20	-0.827204		
ENSECAG00000013998	SQLE					Down18	-0.853104		
ENSECAG00000018059	SC5D					Down16	-0.913739		
ENSECAG00000002626	GPR12					Down10	-0.983593		
ENSECAG00000007020	C14orf28					Down8	-1.07536		
ENSECAG00000011527	CASD1					Down7	-1.1038		
ENSECAG00000021607	TXNDC8					Down6	-1.12117		
ENSECAG00000024444	CTNNA2					Down5	-1.1567		
ENSECAG00000003658	ENSECAG00000003658					Down4	-1.16691		
ENSECAG00000006174	ENSECAG00000006174					Down1	-1.46428		
ENSECAG00000027451	ENSECAG00000027451							Up2	inf
ENSECAG00000016985	IL36G							Up20	1.39083
ENSECAG00000027624	ENSECAG00000027624							Up26	1.18142

APPENDIX 45: List of DEGs IN O/03-infected samples – Part 11

		O/03-WT		O/03-K186E		O/03-230		O/03-K186E-230	
Gene ID	**Gene name**	**Rank**	**Log2FC**	**Rank**	**Log2FC**	**Rank**	**Log2FC**	**Rank**	**Log2FC**
ENSECAG00000015449	IRF8							Up32	1.07107
ENSECAG00000014398	ENSECAG00000014398							Up35	1.01109
ENSECAG00000015389	SV2B							Up38	0.908518
ENSECAG00000022387	RUNDC3B							Up43	0.82593
ENSECAG00000024909	FBXL2							Up54	0.713893
ENSECAG00000015959	MYBL1							Up55	0.705196
ENSECAG00000009742	S100A12							Up57	0.704123
ENSECAG00000024016	IL13RA2							Up61	0.68763
ENSECAG00000011226	CENPQ							Up63	0.6774
ENSECAG00000024391	CD58							Up65	0.667648
ENSECAG00000011292	MSR1							Up70	0.653605
ENSECAG00000009925	ARNTL2							Up71	0.648288
ENSECAG00000023416	F3							Up73	0.64621
ENSECAG00000019890	VRK1							Up74	0.635457
ENSECAG00000020270	CCNE2							Up77	0.631894
ENSECAG00000010871	TMEM126B							Up80	0.623039
ENSECAG00000027692	COX2							Up85	0.611617
ENSECAG00000022189	PFDN4							Up86	0.611601
ENSECAG00000019398	TFPI2							Up88	0.60962
ENSECAG00000019817	SNRPF							Up90	0.608546
ENSECAG00000027681	ND2							Up97	0.597166
ENSECAG00000020878	ENSECAG00000020878							Up101	0.581476
ENSECAG00000017926	TROAP							Down57	-0.582251
ENSECAG00000002482	PQLC1							Down56	-0.58856
ENSECAG00000011939	CLIP2							Down53	-0.595107
ENSECAG00000008012	CHST3							Down52	-0.598346
ENSECAG00000007907	ENSECAG00000007907							Down46	-0.619994
ENSECAG00000009352	OBSL1							Down45	-0.621182
ENSECAG00000017294	SNTB1							Down44	-0.624512
ENSECAG00000021037	FEZF2							Down41	-0.64712
ENSECAG00000017142	SEZ6							Down39	-0.663987
ENSECAG00000000744	TBC1D2							Down34	-0.673914
ENSECAG00000026973	BCL11B							Down21	-0.736266
ENSECAG00000014357	DLL1							Down19	-0.739284
ENSECAG00000008871	ALDH1A3							Down16	-0.75701
ENSECAG00000009779	SORCS2							Down15	-0.772107
ENSECAG00000013825	NOD1							Down14	-0.79745
ENSECAG00000017019	ARHGAP33							Down13	-0.810786
ENSECAG00000011401	CATSPERD							Down12	-0.814234
ENSECAG00000014632	CPLX2							Down10	-0.904945
ENSECAG00000023032	AZIN2							Down6	-1.00068
ENSECAG00000017175	REM1							Down4	-1.09741
ENSECAG00000013783	ENSECAG00000013783							Down3	-1.14067
ENSECAG00000026054	U6							Down1	-inf

		O/03-WT	O/03-K186E	O/03-230	O/03-K186E 230
	Up-regulated genes	248	65	179	101
	Down-regulated genes	181	128	62	57
	Total DEGs	429	193	241	158

APPENDIX 46: List of DEGs IN U/63-infected samples – Part 1

		U/63-WT		U/63-E186K		U/63-219		U/63-E186K-219	
Gene ID	**Gene name**	**Rank**	**Log2FC**	**Rank**	**Log2FC**	**Rank**	**Log2FC**	**Rank**	**Log2FC**
ENSECAG00000010153	IFIT3	Up6	5.33721	Up9	3.44747	Up4	5.80104	Up4	3.81311
ENSECAG00000011439	DHX58	Up9	4.3019	Up11	2.68535	Up5	5.21961	Up7	3.38991
ENSECAG00000009020	IFI44	Up10	4.22229	Up12	2.65132	Up9	4.58574	Up8	3.38161
ENSECAG00000004433	IFIT1	Up11	4.11074	Up14	2.06929	Up7	4.81752	Up10	2.91267
ENSECAG00000009618	IFI44L	Up12	3.63587	Up15	2.05809	Up10	3.90641	Up11	2.72257
ENSECAG00000011776	MX1	Up14	3.14878	Up33	1.28711	Up11	3.69529	Up17	1.89618
ENSECAG00000001324	ISG15	Up18	2.77236	Up24	1.51271	Up12	3.63029	Up18	1.8953
ENSECAG00000012132	ENSECAG00000012132	Up22	2.14977	Up73	0.827483	Up17	2.7074	Up33	1.40367
ENSECAG00000014422	OAS2	Up27	1.82461	Up57	0.941087	Up20	2.34113	Up53	1.2398
ENSECAG00000024429	IRF9	Up28	1.79513	Up44	1.06099	Up19	2.59713	Up43	1.30893
ENSECAG00000001481	SAMD9L	Up32	1.49784	Up123	0.645397	Up24	1.70405	Up52	1.259
ENSECAG00000001598	IFI35	Up48	1.05837	Up70	0.850989	Up26	1.62154	Up106	1.00694
ENSECAG00000023135	LYRM7	Up64	0.928043	Up85	0.74354	Up76	0.718885	Up114	0.995131
ENSECAG00000012179	RGS2	Up89	0.751255	Up36	1.19713	Up85	0.665814	Up57	1.2091
ENSECAG00000015959	MYBL1	Up133	0.595256	Up99	0.710276	Up108	0.604595	Up140	0.966431
ENSECAG00000015814	LRRC73	Down39	-0.90033	Down37	-0.675876	Down81	-0.63939	Down53	-0.618326
ENSECAG00000009202	TYSND1	Down23	-0.959203	Down55	-0.597133	Down88	-0.621765	Down46	-0.664262
ENSECAG00000008899	LSS	Down22	-0.963068	Down28	-0.745571	Down65	-0.726256	Down20	-0.80581
ENSECAG00000010664	GNRH1	Up4	inf	Up2	inf			Up2	inf
ENSECAG00000020860	LSMEM1	Up15	3.00232	Up7	3.52154			Up14	2.54166
ENSECAG00000009793	TAGLN3	Up20	2.54268	Up8	3.49492			Up9	2.98379
ENSECAG00000019367	ENSECAG00000019367	Up25	1.88129	Up16	1.96328			Up16	1.97533
ENSECAG00000022299	LRRC66	Up26	1.82795	Up13	2.21072			Up19	1.8297
ENSECAG00000022605	FAM179A	Up29	1.74385	Up20	1.62461			Up129	0.978574
ENSECAG00000012898	PNLDC1	Up31	1.56451	Up23	1.55336			Up92	1.05295
ENSECAG00000020923	LEAP2	Up37	1.29596	Up27	1.42569			Up44	1.30241
ENSECAG00000012348	EFNB2	Up39	1.27502	Up35	1.21889			Up34	1.3969
ENSECAG00000023507	GPR19	Up42	1.15431	Up40	1.10944			Up237	0.849092
ENSECAG00000016648	IL5	Up44	1.13772	Up39	1.11771			Up99	1.03268
ENSECAG00000010656	HBEGF	Up52	1.01248	Up22	1.58423			Up21	1.78032
ENSECAG00000015342	CXCL8	Up61	0.956865	Up58	0.935643			Up25	1.59604
ENSECAG00000023566	DHRS9	Up75	0.853005	Up56	0.944545			Up26	1.57134
ENSECAG00000005903	ENSECAG00000005903	Up76	0.849148	Up77	0.806524			Up153	0.949382
ENSECAG00000024298	PTPN22	Up78	0.824954	Up52	1.00397			Up155	0.93912
ENSECAG00000013756	C8orf59	Up84	0.783365	Up120	0.650365			Up310	0.80989
ENSECAG00000020270	CCNE2	Up88	0.758335	Up48	1.01197			Up84	1.06706
ENSECAG00000022874	SCG5	Up93	0.737912	Up63	0.891699			Up567	0.705102
ENSECAG00000016965	SUV39H2	Up94	0.736519	Up64	0.886631			Up90	1.05648
ENSECAG00000013843	TUBE1	Up97	0.730053	Up69	0.853153			Up45	1.29863
ENSECAG00000023484	SKA1	Up101	0.706982	Up107	0.691046			Up277	0.829442
ENSECAG00000007803	LSM3	Up102	0.704097	Up109	0.677778			Up108	1.00363
ENSECAG00000009925	ARNTL2	Up103	0.703787	Up67	0.862834			Up79	1.08686
ENSECAG00000015496	ENSECAG00000015496	Up104	0.702803	Up75	0.819189			Up294	0.818364
ENSECAG00000003755	PIGW	Up105	0.701986	Up80	0.788466			Up228	0.8568
ENSECAG00000015510	FGF7	Up106	0.695944	Up45	1.02384			Up31	1.48791
ENSECAG00000004590	TIPIN	Up109	0.676348	Up74	0.821189			Up274	0.830524
ENSECAG00000014615	STX2	Up110	0.672309	Up134	0.611084			Up145	0.959706
ENSECAG00000017194	FANCB	Up111	0.670598	Up118	0.656224			Up136	0.969754
ENSECAG00000020201	GCNT1	Up114	0.661875	Up110	0.672867			Up967	0.600993
ENSECAG00000013298	POLE2	Up118	0.648491	Up121	0.647938			Up102	1.02063
ENSECAG00000024391	CD58	Up122	0.639183	Up88	0.735449			Up50	1.26502
ENSECAG00000011449	ANKS1B	Up123	0.638934	Up55	0.965427			Up477	0.736421
ENSECAG00000007070	TPRKB	Up126	0.625957	Up92	0.730267			Up113	0.996073
ENSECAG00000002238	MIS12	Up130	0.60541	Up82	0.780604			Up292	0.81948
ENSECAG00000009754	NETO2	Up134	0.594759	Up144	0.589724			Up376	0.775708
ENSECAG00000006472	EIF4E	Up135	0.590533	Up145	0.589428			Up96	1.04404
ENSECAG00000023637	ADAM33	Down235	-0.593988	Down29	-0.742143			Down27	-0.733496
ENSECAG00000005577	GPR39	Down114	-0.697544	Down27	-0.76466			Down47	-0.662203
ENSECAG00000001989	SESN2	Down46	-0.879919	Down21	-0.833335			Down64	-0.585299
ENSECAG00000009989	CHAC1	Down18	-0.993054	Down16	-0.905742			Down37	-0.691439
ENSECAG00000021543	CDKN1A	Down10	-1.09995	Down54	-0.597673			Down58	-0.60167
ENSECAG00000007881	IFIH1	Up7	4.55421			Up6	5.13002	Up5	3.65116
ENSECAG00000024167	DDX60	Up8	4.43945			Up8	4.69968	Up6	3.4278
ENSECAG00000000284	ENSECAG00000000284	Up19	2.72162			Up14	3.36723	Up28	1.54525
ENSECAG00000010185	ENSECAG00000010185	Up21	2.3283			Up16	3.03826	Up29	1.53936
ENSECAG00000019932	ENSECAG00000019932	Up23	1.9909			Up18	2.64499	Up27	1.56487
ENSECAG00000000968	ENSECAG00000000968	Up30	1.613			Up23	1.83489	Up116	0.992858
ENSECAG00000021220	ENSECAG00000021220	Up38	1.29564			Up21	2.04664	Up373	0.776238
ENSECAG00000008809	OAS3	Up40	1.2671			Up22	1.93784	Up473	0.740666
ENSECAG00000011726	EIF2AK2	Up49	1.04951			Up50	0.944299	Up120	0.988205
ENSECAG00000021989	DDX58	Up50	1.03753			Up35	1.34299	Up563	0.705945
ENSECAG00000007744	RSRP1	Up53	1.01076			Up47	1.00543	Up35	1.38565
ENSECAG00000009686	DTX3L	Up55	1.00164			Up28	1.46786	Up437	0.755732

APPENDIX 47: List of DEGs IN U/63-infected samples – Part 2

		U/63-WT		U/63-E186K		U/63-219		U/63-E186K-219	
Gene ID	Gene name	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC
ENSECAG00000012331	PARP9	Up56	0.998805			Up34	1.34397	Up296	0.816214
ENSECAG00000016217	RNF213	Up65	0.923765			Up27	1.49935	Up705	0.663515
ENSECAG00000001399	SAMD9	Up66	0.912912			Up52	0.942535	Up211	0.870057
ENSECAG00000017394	ZC3HAV1	Up68	0.903109			Up37	1.30002	Up959	0.602414
ENSECAG00000004349	IFIT5	Up71	0.885287			Up48	0.985357	Up845	0.625025
ENSECAG00000009543	EPSTI1	Up77	0.83257			Up29	1.4297	Up279	0.828657
ENSECAG00000014295	PI3	Up83	0.785474			Up32	1.37746	Up217	0.86702
ENSECAG00000016983	ANKRD1	Up85	0.780023			Up40	1.09562	Up259	0.838183
ENSECAG00000020428	ENSECAG00000020428	Up116	0.657532			Up64	0.793587	Up223	0.861115
ENSECAG00000010667	ERG	Up128	0.610224			Up71	0.747632	Up551	0.709218
ENSECAG00000008688	ENSECAG00000008688	Down189	-0.625213			Down3	-1.77432	Down59	-0.599041
ENSECAG00000024650	FADS2	Down181	-0.63414			Down52	-0.799129	Down41	-0.670507
ENSECAG00000013009	KIF26B	Down138	-0.677023			Down77	-0.649226	Down24	-0.737876
ENSECAG00000013608	KAZALD1	Down123	-0.687527			Down54	-0.783873	Down54	-0.614904
ENSECAG00000021964	ASB9	Down117	-0.695657			Down33	-0.916693	Down32	-0.71169
ENSECAG00000023067	MVK	Down96	-0.717135			Down84	-0.627694	Down38	-0.687078
ENSECAG00000018552	CDKN2A	Down64	-0.800677			Down71	-0.692934	Down45	-0.665138
ENSECAG00000016389	ACSS2	Down50	-0.860334			Down56	-0.770018	Down19	-0.824737
ENSECAG00000017229	ENSECAG00000017229	Down34	-0.907961			Down13	-1.16944	Down49	-0.637804
ENSECAG00000006999	PLEKHG4	Down33	-0.909293			Down72	-0.686726	Down18	-0.838619
ENSECAG00000013176	MEGF6	Down26	-0.939986			Down39	-0.877218	Down52	-0.63014
ENSECAG00000005316	PIF1	Down25	-0.940866			Down37	-0.906174	Down13	-0.901346
ENSECAG00000013300	CXCL10	Up13	3.61765					Up12	2.59049
ENSECAG00000003681	ENSECAG00000003681	Up41	1.19309					Up152	0.952266
ENSECAG00000005249	IQUB	Up45	1.12424					Up55	1.21895
ENSECAG00000020285	NMUR2	Up59	0.964004					Up314	0.808767
ENSECAG00000015389	SV2B	Up72	0.880344					Up225	0.858517
ENSECAG00000012680	N4BP2L1	Up73	0.877394					Up172	0.908449
ENSECAG00000019165	CENPK	Up90	0.748959					Up75	1.11258
ENSECAG00000022568	TMEM154	Up95	0.732243					Up770	0.648252
ENSECAG00000008723	CASP12	Up96	0.73034					Up268	0.832213
ENSECAG00000016844	MEIOC	Up98	0.720769					Up412	0.762811
ENSECAG00000023416	F3	Up99	0.716922					Up171	0.913206
ENSECAG00000016730	ENSECAG00000016730	Up100	0.713195					Up163	0.926146
ENSECAG00000018861	ENSECAG00000018861	Up112	0.670498					Up267	0.832327
ENSECAG00000005331	ISOC1	Up113	0.662094					Up288	0.822006
ENSECAG00000024519	SPC25	Up115	0.658804					Up744	0.654211
ENSECAG00000024911	ITGB3BP	Up119	0.647272					Up135	0.970525
ENSECAG00000014972	PHF11	Up120	0.641586					Up443	0.752192
ENSECAG00000023743	CCDC58	Up124	0.637457					Up585	0.696651
ENSECAG00000018693	ENSECAG00000018693	Up127	0.621964					Up989	0.595197
ENSECAG00000003015	SERPINB2	Up131	0.604312					Up530	0.717132
ENSECAG00000018544	SPDL1	Up136	0.589304					Up198	0.884021
ENSECAG00000019024	FAM204A	Up137	0.58901					Up713	0.661039
ENSECAG00000009006	USP46	Up138	0.58693					Up272	0.830979
ENSECAG00000019053	FAM175A	Up140	0.581694					Up74	1.12084
ENSECAG00000010225	GRID2IP	Down165	-0.647613					Down25	-0.735195
ENSECAG00000021280	SHISA4	Down164	-0.647663					Down60	-0.595579
ENSECAG00000022901	PKMYT1	Down152	-0.658362					Down57	-0.601905
ENSECAG00000003075	BORCS6	Down98	-0.712612					Down23	-0.751196
ENSECAG00000011160	TP53INP2	Down82	-0.741446					Down31	-0.715738
ENSECAG00000010118	SRGAP3	Down80	-0.747885					Down43	-0.667821
ENSECAG00000000931	TMEM119	Down79	-0.748376					Down30	-0.718741
ENSECAG00000021186	FHOD3	Down71	-0.771865					Down61	-0.590052
ENSECAG00000021155	MXD3	Down69	-0.773135					Down10	-1.01078
ENSECAG00000004462	HIC1	Down66	-0.794084					Down63	-0.586445
ENSECAG00000008443	TMEM259	Down54	-0.835336					Down51	-0.637604
ENSECAG00000014632	CPLX2	Down52	-0.838319					Down33	-0.710916
ENSECAG00000011206	MVD	Down51	-0.859346					Down56	-0.611066
ENSECAG00000019910	PDLIM2	Down40	-0.899907					Down28	-0.73035
ENSECAG00000012197	SPATA2L	Down38	-0.900881					Down35	-0.701257
ENSECAG00000001022	UCN2	Down36	-0.905955					Down34	-0.704328
ENSECAG00000021411	TM7SF2	Down32	-0.913836					Down22	-0.752184
ENSECAG00000011185	BTBD19	Down15	-1.0045					Down21	-0.786424
ENSECAG00000023782	SLC16A13	Down12	-1.0319					Down9	-1.01198
ENSECAG00000014282	NPPB	Down4	-1.37353					Down7	-1.06006
ENSECAG00000010322	LDHD	Down3	-1.39661					Down4	-1.20231
ENSECAG00000006098	ANKRD9	Down1	-1.99072					Down1	-2.86832
ENSECAG00000025973	U1			Up6	inf	Up3	inf		
ENSECAG00000008637	HMGCS1			Down56	-0.594789	Down68	-0.710506		
ENSECAG00000008871	ALDH1A3			Down4	-1.13855	Up90	0.652271		
ENSECAG00000018139	ENSECAG00000018139	Up1	inf	Up3	inf				
ENSECAG00000019576	ENSECAG00000019576	Up2	inf	Up4	inf				

APPENDIX 48: List of DEGs IN U/63-infected samples – Part 3

		U/63-WT		U/63-E186K		U/63-219		U/63-E186K-219	
Gene ID	Gene name	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC
ENSECAG00000022734	SGPP2	Up3	inf	Up5	inf				
ENSECAG00000021468	POPDC2	Up34	1.34092	Up21	1.59069				
ENSECAG00000016080	KCNB1	Up35	1.33649	Up30	1.32885				
ENSECAG00000012746	STAR	Up36	1.31773	Up28	1.39059				
ENSECAG00000012826	CDO1	Up43	1.14266	Up26	1.45509				
ENSECAG00000018029	THSD1	Up54	1.00861	Up19	1.69768				
ENSECAG00000013528	AVIL	Up57	0.972598	Up83	0.754989				
ENSECAG00000020562	AXDND1	Up69	0.890887	Up38	1.13744				
ENSECAG00000003669	HLA-DMA	Up70	0.888328	Up32	1.29289				
ENSECAG00000026993	CEBPZOS	Up80	0.81461	Up96	0.722319				
ENSECAG00000023052	POLG2	Up81	0.813714	Up105	0.696331				
ENSECAG00000009937	UBE2N	Up82	0.805787	Up116	0.659094				
ENSECAG00000001215	C12orf73	Up86	0.768625	Up129	0.618742				
ENSECAG00000010871	TMEM126B	Up129	0.607189	Up124	0.63787				
ENSECAG00000011939	CLIP2	Down214	-0.60504	Down34	-0.71819				
ENSECAG00000011401	CATSPERD	Down209	-0.606865	Down9	-1.00675				
ENSECAG00000014566	WISP2	Down206	-0.608505	Down41	-0.66546				
ENSECAG00000014470	EEPD1	Down194	-0.618182	Down22	-0.819937				
ENSECAG00000009779	SORCS2	Down190	-0.624447	Down11	-0.97256				
ENSECAG00000008298	WNT10B	Down173	-0.640325	Down53	-0.59968				
ENSECAG00000013825	NOD1	Down166	-0.646151	Down30	-0.739632				
ENSECAG00000009792	UNC5B	Down143	-0.668636	Down7	-1.05305				
ENSECAG00000005067	KRT16	Down118	-0.693857	Down3	-1.14763				
ENSECAG00000012059	OPLAH	Down106	-0.707164	Down20	-0.840835				
ENSECAG00000012961	HSPG2	Down103	-0.708486	Down58	-0.585461				
ENSECAG00000007907	ENSECAG00000007907	Down83	-0.74087	Down36	-0.682479				
ENSECAG00000005017	FBXW9	Down74	-0.762503	Down23	-0.804287				
ENSECAG00000013443	AKNA	Down56	-0.823496	Down44	-0.633176				
ENSECAG00000000607	TRIB3	Down29	-0.923731	Down24	-0.793802				
ENSECAG00000013817	CYP1A1	Down28	-0.928951	Down1	-2.00177				
ENSECAG00000011234	COL7A1	Down27	-0.933107	Down45	-0.621001				
ENSECAG00000011475	ENSECAG00000011475	Down9	-1.10002	Down42	-0.661868				
ENSECAG00000010953	TNIP3					Up54	0.932075	Up890	0.616436
ENSECAG00000019809	MB21D1					Up72	0.745114	Up442	0.752859
ENSECAG00000017105	SEC62					Up83	0.671821	Up954	0.603615
ENSECAG00000008426	PAPSS2					Up87	0.66224	Up538	0.714752
ENSECAG00000009472	CLSPN					Up95	0.629392	Up1020	0.589565
ENSECAG00000019817	SNRPF					Up96	0.629116	Up769	0.648327
ENSECAG00000022166	AKAP12					Up98	0.626063	Up493	0.730739
ENSECAG00000003783	ENSECAG00000003783					Up102	0.613639	Up965	0.601284
ENSECAG00000013778	PINX1					Up113	0.592925	Up859	0.622534
ENSECAG00000015715	ADAMTS15					Down24	-1.01545	Down39	-0.682253
ENSECAG00000024260	TMEM160					Down21	-1.04154	Down29	-0.7263
ENSECAG00000020157	KRT5					Down11	-1.23713	Down15	-0.893729
ENSECAG00000003332	H1FX					Down8	-1.30571	Down11	-0.947437
ENSECAG00000027624	ENSECAG00000027624					Down2	-2.0526	Down3	-1.21182
ENSECAG00000024888	CCL5	Up16	2.79473			Up13	3.37828		
ENSECAG00000019309	GBP2	Up17	2.78832			Up15	3.33191		
ENSECAG00000019411	HERC6	Up46	1.12094			Up38	1.29109		
ENSECAG00000002176	ENSECAG00000002176	Up51	1.01873			Up25	1.62778		
ENSECAG00000003585	RPL19	Up67	0.907307			Up31	1.41658		
ENSECAG00000010055	RSAD2	Up74	0.856092			Up36	1.31333		
ENSECAG00000004574	ENSECAG00000004574	Up87	0.760706			Up49	0.967775		
ENSECAG00000012773	CMPK2	Up91	0.74871			Up46	1.02271		
ENSECAG00000009190	TMSB4X	Up108	0.678294			Up61	0.829304		
ENSECAG00000013435	OAS1	Up117	0.655917			Up41	1.08034		
ENSECAG00000024399	FGF12	Up139	0.583093			Up88	0.6566		
ENSECAG00000010262	PERP	Down247	-0.586801			Down87	-0.623274		
ENSECAG00000012818	TUBB3	Down215	-0.6047			Down86	-0.624004		
ENSECAG00000013744	HSPB6	Down204	-0.610737			Down85	-0.626631		
ENSECAG00000005956	TPBGL	Down183	-0.630657			Down64	-0.728843		
ENSECAG00000006734	ENSECAG00000006734	Down146	-0.66453			Down63	-0.734948		
ENSECAG00000021748	ENSECAG00000021748	Down130	-0.683528			Down58	-0.7599		
ENSECAG00000021308	APBA3	Down105	-0.707515			Down42	-0.865827		
ENSECAG00000015240	PROM2	Down101	-0.709566			Down99	-0.58243		
ENSECAG00000020198	NOTUM	Down91	-0.728727			Down23	-1.02528		
ENSECAG00000016441	DOHH	Down87	-0.734432			Down80	-0.640043		
ENSECAG00000010652	ENSECAG00000010652	Down73	-0.762938			Down35	-0.91315		
ENSECAG00000024586	ENSECAG00000024586	Down31	-0.913975			Down40	-0.873208		
ENSECAG00000025020	FHL1	Down21	-0.981485			Down15	-1.12762		
ENSECAG00000018561	CDKN2B	Down14	-1.01255			Down30	-0.940305		
ENSECAG00000024000	PYGM			Up10	3.2163			Up13	2.58043
ENSECAG00000023733	MMP1			Up17	1.7786			Up15	2.11539

APPENDIX 49: List of DEGs IN U/63-infected samples – Part 4

		U/63-WT		U/63-E186K		U/63-219		U/63-E186K-219	
Gene ID	Gene name	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC
ENSECAG00000010700	CREG2			Up29	1.36946			Up207	0.876578
ENSECAG00000011626	S1PR1			Up31	1.30539			Up58	1.19872
ENSECAG00000012553	HSF2			Up37	1.18814			Up809	0.635701
ENSECAG00000006567	SLITRK4			Up41	1.10058			Up54	1.22945
ENSECAG00000007918	TNS4			Up43	1.06618			Up740	0.655654
ENSECAG00000023729	SHC3			Up47	1.01967			Up587	0.695941
ENSECAG00000020145	FBXO43			Up51	1.00485			Up93	1.05037
ENSECAG00000008533	ITGA10			Up54	0.976869			Up453	0.750406
ENSECAG00000019438	OSMR			Up59	0.932801			Up440	0.753755
ENSECAG00000013106	ESM1			Up61	0.908903			Up146	0.959285
ENSECAG00000000152	KITLG			Up62	0.905236			Up65	1.14341
ENSECAG00000021206	HAS2			Up68	0.85518			Up97	1.03863
ENSECAG00000020380	COMMD8			Up71	0.845789			Up72	1.12474
ENSECAG00000011003	PDE4B			Up72	0.841612			Up688	0.666768
ENSECAG00000013514	DUSP5			Up78	0.799917			Up142	0.963009
ENSECAG00000011978	LRR1			Up79	0.79051			Up158	0.934407
ENSECAG00000016457	E2F8			Up86	0.741605			Up807	0.636158
ENSECAG00000016339	ADAMTS1			Up87	0.737078			Up372	0.776367
ENSECAG00000011630	FBXO5			Up90	0.734096			Up89	1.05982
ENSECAG00000018894	RPL22L1			Up91	0.730381			Up105	1.01037
ENSECAG00000014572	ENSECAG00000014572			Up93	0.725896			Up233	0.851117
ENSECAG00000011186	ENSECAG00000011186			Up94	0.72516			Up982	0.597596
ENSECAG00000019989	SNRPG			Up95	0.724849			Up182	0.896792
ENSECAG00000002414	SGTB			Up97	0.717154			Up69	1.13378
ENSECAG00000020295	GEN1			Up98	0.714695			Up188	0.889261
ENSECAG00000012899	ESCO2			Up100	0.707083			Up115	0.993846
ENSECAG00000002250	SPRY2			Up101	0.705795			Up245	0.843575
ENSECAG00000011364	RNF219			Up102	0.702664			Up124	0.985682
ENSECAG00000004120	ABHD18			Up104	0.700428			Up164	0.925574
ENSECAG00000011226	CENPQ			Up108	0.683434			Up250	0.84234
ENSECAG00000021234	NET1			Up111	0.671168			Up457	0.749431
ENSECAG00000017783	FST			Up112	0.665451			Up759	0.650341
ENSECAG00000026862	GCSH			Up113	0.664199			Up149	0.955276
ENSECAG00000018634	UBASH3B			Up114	0.6607			Up494	0.729732
ENSECAG00000015790	SLF1			Up115	0.660057			Up68	1.1352
ENSECAG00000007885	ORC5			Up117	0.657888			Up271	0.83168
ENSECAG00000022189	PFDN4			Up119	0.650603			Up47	1.28856
ENSECAG00000015003	PLGRKT			Up122	0.646453			Up111	0.997463
ENSECAG00000022137	DNA2			Up125	0.635694			Up239	0.848146
ENSECAG00000022964	CDC7			Up126	0.634887			Up224	0.860216
ENSECAG00000024016	IL13RA2			Up127	0.630485			Up71	1.12539
ENSECAG00000002211	CD80			Up128	0.623285			Up232	0.852307
ENSECAG00000026888	MRPL42			Up130	0.616			Up1031	0.58653
ENSECAG00000018443	ATAD5			Up131	0.614103			Up213	0.869753
ENSECAG00000009576	MMS22L			Up132	0.612656			Up179	0.901609
ENSECAG00000002513	ERCC6L			Up133	0.611934			Up367	0.777908
ENSECAG00000018810	HELLS			Up135	0.610723			Up327	0.804216
ENSECAG00000009270	ORC2			Up136	0.603078			Up201	0.883252
ENSECAG00000009716	CDK1			Up137	0.602965			Up218	0.866389
ENSECAG00000015376	GLRX2			Up138	0.596			Up821	0.632132
ENSECAG00000022495	NKRF			Up139	0.593306			Up407	0.766519
ENSECAG00000017572	ZNF280C			Up141	0.592528			Up94	1.0462
ENSECAG00000008852	ABCA1			Up142	0.592405			Up561	0.706499
ENSECAG00000022190	GALNT1			Up143	0.591049			Up125	0.983941
ENSECAG00000016959	B3GNT2			Up146	0.588384			Up752	0.652061
ENSECAG00000003773	MICALCL			Up53	1.00184	Up43	1.05852	Up64	1.14546
ENSECAG00000026524	U2	Up5	inf						
ENSECAG00000003230	AKAP5	Up24	1.88967						
ENSECAG00000015717	TSSK3	Up33	1.48859						
ENSECAG00000023835	KLHL32	Up47	1.06422						
ENSECAG00000014398	ENSECAG00000014398	Up58	0.967904						
ENSECAG00000012653	DYRK4	Up60	0.963909						
ENSECAG00000012801	KBTBD3	Up62	0.940261						
ENSECAG00000015383	ITGB1BP2	Up63	0.935973						
ENSECAG00000026834	TMEM167A	Up79	0.8229						
ENSECAG00000026955	SELK	Up92	0.741346						
ENSECAG00000023753	OGFRL1	Up107	0.694384						
ENSECAG00000001334	C3orf14	Up121	0.640623						
ENSECAG00000021036	LSM6	Up125	0.627901						
ENSECAG00000026824	LYRM4	Up132	0.595951						
ENSECAG00000008435	LMF2	Down258	-0.582244						
ENSECAG00000012006	BCL9L	Down257	-0.582536						
ENSECAG00000009361	COL5A1	Down256	-0.583028						

APPENDIX 50: List of DEGs IN U/63-infected samples – Part 5

		U/63-WT		U/63-E186K		U/63-219		U/63-E186K-219	
Gene ID	Gene name	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC
ENSECAG00000019381	MEGF8	Down255	-0.583606						
ENSECAG00000015071	PTRF	Down254	-0.584623						
ENSECAG00000019242	PLXNB2	Down253	-0.584943						
ENSECAG00000019803	GALK1	Down252	-0.585071						
ENSECAG00000010237	FBXL6	Down251	-0.58569						
ENSECAG00000013823	SCRIB	Down250	-0.586134						
ENSECAG00000022794	ENSECAG00000022794	Down249	-0.586204						
ENSECAG00000025129	NDRG1	Down248	-0.58647						
ENSECAG00000024665	GHDC	Down246	-0.587181						
ENSECAG00000019465	PLTP	Down245	-0.587261						
ENSECAG00000000369	CERS4	Down244	-0.58816						
ENSECAG00000004871	SERPINH1	Down243	-0.588866						
ENSECAG00000007152	ENSECAG00000007152	Down242	-0.590067						
ENSECAG00000009826	SCRN2	Down241	-0.590353						
ENSECAG00000014242	MAD1L1	Down240	-0.591694						
ENSECAG00000014699	NUAK2	Down239	-0.5923						
ENSECAG00000024119	COL16A1	Down238	-0.592302						
ENSECAG00000010619	F12	Down237	-0.592422						
ENSECAG00000005169	ABCC10	Down236	-0.592478						
ENSECAG00000007647	COL5A3	Down234	-0.595539						
ENSECAG00000007338	SCARF2	Down233	-0.597139						
ENSECAG00000024863	ENSECAG00000024863	Down232	-0.597405						
ENSECAG00000002876	FAAP100	Down231	-0.597518						
ENSECAG00000019788	WDR83	Down230	-0.597677						
ENSECAG00000018063	D2HGDH	Down229	-0.598239						
ENSECAG00000024966	RAVER1	Down228	-0.598494						
ENSECAG00000021805	LAMB2	Down227	-0.599262						
ENSECAG00000007334	SMG9	Down226	-0.599316						
ENSECAG00000005228	PHGDH	Down225	-0.600207						
ENSECAG00000019299	TBL3	Down224	-0.600235						
ENSECAG00000000797	SUN2	Down223	-0.600716						
ENSECAG00000000402	NDUFS7	Down222	-0.601934						
ENSECAG00000012599	TRMT61A	Down221	-0.602192						
ENSECAG00000026879	TXNRD2	Down220	-0.60288						
ENSECAG00000020780	PELP1	Down219	-0.603108						
ENSECAG00000007474	SPHK2	Down218	-0.603315						
ENSECAG00000016486	ARRDC2	Down217	-0.603381						
ENSECAG00000021290	MBD6	Down216	-0.603718						
ENSECAG00000015054	XXYLT1	Down213	-0.605379						
ENSECAG00000025091	MVP	Down212	-0.606154						
ENSECAG00000015206	SYNE3	Down211	-0.606585						
ENSECAG00000016933	FZR1	Down210	-0.606742						
ENSECAG00000007611	INF2	Down208	-0.607694						
ENSECAG00000012180	NAGPA	Down207	-0.607697						
ENSECAG00000013037	ALDOC	Down205	-0.609886						
ENSECAG00000015488	KCNN4	Down203	-0.612627						
ENSECAG00000002668	RHOB	Down202	-0.614745						
ENSECAG00000003436	APBB3	Down201	-0.615551						
ENSECAG00000000744	TBC1D2	Down200	-0.616395						
ENSECAG00000010992	NFATC4	Down199	-0.616504						
ENSECAG00000017794	IRF1	Down198	-0.616974						
ENSECAG00000011852	ALDH4A1	Down197	-0.617007						
ENSECAG00000007272	NCOR2	Down196	-0.617372						
ENSECAG00000022228	KLC4	Down195	-0.617479						
ENSECAG00000009550	FLNA	Down193	-0.619276						
ENSECAG00000011133	HSD3B7	Down192	-0.62008						
ENSECAG00000017863	PFKL	Down191	-0.623694						
ENSECAG00000020716	MYH9	Down188	-0.625961						
ENSECAG00000024224	WDR34	Down187	-0.626388						
ENSECAG00000023595	ENSECAG00000023595	Down186	-0.628599						
ENSECAG00000024038	YPEL3	Down185	-0.629591						
ENSECAG00000006063	JRK	Down184	-0.630245						
ENSECAG00000021765	TSPO	Down182	-0.632928						
ENSECAG00000006963	NACC2	Down180	-0.634188						
ENSECAG00000017371	SMTN	Down179	-0.636204						
ENSECAG00000016064	SYDE1	Down178	-0.636416						
ENSECAG00000021059	MED16	Down177	-0.637326						
ENSECAG00000021785	MHCB3	Down176	-0.637447						
ENSECAG00000021774	DNAAF5	Down175	-0.637672						
ENSECAG00000023341	DHCR7	Down174	-0.638126						
ENSECAG00000018860	ZFP36L1	Down172	-0.640812						
ENSECAG00000004107	DDIT4	Down171	-0.641177						
ENSECAG00000014491	NLRC5	Down170	-0.641394						

APPENDIX 51: List of DEGs IN U/63-infected samples – Part 6

		U/63-WT		U/63-E186K		U/63-219		U/63-E186K-219	
Gene ID	Gene name	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC
ENSECAG00000017998	DHCR24	Down169	-0.644519						
ENSECAG00000015944	TKFC	Down168	-0.644629						
ENSECAG00000021659	CCDC157	Down167	-0.645226						
ENSECAG00000009314	TMEM132A	Down163	-0.647746						
ENSECAG00000015936	ENSECAG00000015936	Down162	-0.648071						
ENSECAG00000012654	EMC10	Down161	-0.649023						
ENSECAG00000006501	FJX1	Down160	-0.650037						
ENSECAG00000012635	BRAT1	Down159	-0.651806						
ENSECAG00000010792	INTS1	Down158	-0.653575						
ENSECAG00000018273	TFPT	Down157	-0.65384						
ENSECAG00000022294	PITX3	Down156	-0.654148						
ENSECAG00000021606	WDR81	Down155	-0.654419						
ENSECAG00000013515	SEMA3B	Down154	-0.657088						
ENSECAG00000011354	DHX34	Down153	-0.657335						
ENSECAG00000012169	ATP5D	Down151	-0.658918						
ENSECAG00000003772	DEDD2	Down150	-0.658949						
ENSECAG00000024481	AHNAK2	Down149	-0.659233						
ENSECAG00000011779	ARID3A	Down148	-0.662581						
ENSECAG00000009818	AMOTL2	Down147	-0.66367						
ENSECAG00000012166	LRP1	Down145	-0.666203						
ENSECAG00000012712	PNPLA6	Down144	-0.668635						
ENSECAG00000024608	SREBF2	Down142	-0.669327						
ENSECAG00000003570	PICK1	Down141	-0.671165						
ENSECAG00000010395	RIN1	Down140	-0.671403						
ENSECAG00000022578	HSPBP1	Down139	-0.673144						
ENSECAG00000018111	NAPRT	Down137	-0.677179						
ENSECAG00000018924	PTK2B	Down136	-0.677781						
ENSECAG00000019633	IGFBP6	Down135	-0.678353						
ENSECAG00000014761	LONP1	Down134	-0.679579						
ENSECAG00000021979	CARD10	Down133	-0.680155						
ENSECAG00000006931	ATF4	Down132	-0.681204						
ENSECAG00000007115	FRMD8	Down131	-0.683335						
ENSECAG00000006908	CD248	Down129	-0.683685						
ENSECAG00000008993	HIP1R	Down128	-0.685781						
ENSECAG00000020178	WNT9A	Down127	-0.685894						
ENSECAG00000016374	ENSECAG00000016374	Down126	-0.686582						
ENSECAG00000011353	MFSD7	Down125	-0.686602						
ENSECAG00000022526	CDH15	Down124	-0.687095						
ENSECAG00000019236	LDLRAD3	Down122	-0.688462						
ENSECAG00000009562	ZC3H12A	Down121	-0.68927						
ENSECAG00000019238	ENSECAG00000019238	Down120	-0.691056						
ENSECAG00000016798	MIEF2	Down119	-0.691551						
ENSECAG00000005097	IFI27	Down116	-0.696728						
ENSECAG00000020730	RNH1	Down115	-0.697501						
ENSECAG00000003836	LRRC75A	Down113	-0.698118						
ENSECAG00000017789	SIX2	Down112	-0.700091						
ENSECAG00000022697	TRABD2B	Down111	-0.700461						
ENSECAG00000019882	CALHM3	Down110	-0.702522						
ENSECAG00000003591	TPGS1	Down109	-0.703566						
ENSECAG00000021641	SNX21	Down108	-0.704026						
ENSECAG00000011011	SNAPC2	Down107	-0.705215						
ENSECAG00000009170	PINK1	Down104	-0.707591						
ENSECAG00000019969	CCDC142	Down102	-0.70912						
ENSECAG00000022949	WFDC3	Down100	-0.709883						
ENSECAG00000006619	ZNF688	Down99	-0.710044						
ENSECAG00000019359	DUS3L	Down97	-0.716211						
ENSECAG00000014320	PPL	Down95	-0.717457						
ENSECAG00000002911	KCTD11	Down94	-0.727666						
ENSECAG00000000706	MEGF11	Down93	-0.727672						
ENSECAG00000002928	FAM212B	Down92	-0.727917						
ENSECAG00000010566	ENSECAG00000010566	Down90	-0.729019						
ENSECAG00000005302	GJB5	Down89	-0.729846						
ENSECAG00000011246	GLI1	Down88	-0.730742						
ENSECAG00000001181	CBX4	Down86	-0.735125						
ENSECAG00000013685	TNXB	Down85	-0.735295						
ENSECAG00000019557	LDLR	Down84	-0.739951						
ENSECAG00000024412	ARHGAP39	Down81	-0.74398						
ENSECAG00000023023	OSGIN1	Down78	-0.755167						
ENSECAG00000015448	TAPBP	Down77	-0.759283						
ENSECAG00000008194	MIDN	Down76	-0.759865						
ENSECAG00000012500	AP5Z1	Down75	-0.761725						
ENSECAG00000011704	TNFAIP2	Down72	-0.769714						
ENSECAG00000007112	ENSECAG00000007112	Down70	-0.772325						

APPENDIX 52: List of DEGs IN U/63-infected samples – Part 7

		U/63-WT		U/63-E186K		U/63-219		U/63-E186K-219	
Gene ID	Gene name	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC
ENSECAG00000023969	ENSECAG00000023969	Down68	-0.774057						
ENSECAG00000020219	ENSECAG00000020219	Down67	-0.780144						
ENSECAG00000012116	CARNS1	Down65	-0.798911						
ENSECAG00000020592	JUP	Down63	-0.80294						
ENSECAG00000011097	LMNA	Down62	-0.80346						
ENSECAG00000026973	BCL11B	Down61	-0.804756						
ENSECAG00000015514	PLEC	Down60	-0.809049						
ENSECAG00000023274	LAMA5	Down59	-0.813921						
ENSECAG00000024458	KCTD7	Down58	-0.814027						
ENSECAG00000017961	ENSECAG00000017961	Down57	-0.814127						
ENSECAG00000002061	NEURL2	Down55	-0.834217						
ENSECAG00000021487	TRIM66	Down53	-0.835356						
ENSECAG00000014794	MIIP	Down49	-0.861415						
ENSECAG00000008566	CTSE	Down48	-0.866538						
ENSECAG00000014899	ENSECAG00000014899	Down47	-0.877504						
ENSECAG00000009375	FAM198A	Down45	-0.890646						
ENSECAG00000019125	FASN	Down44	-0.892033						
ENSECAG00000004572	CH25H	Down43	-0.895983						
ENSECAG00000016720	OBSCN	Down42	-0.897825						
ENSECAG00000022258	SDSL	Down41	-0.899809						
ENSECAG00000020167	RGMA	Down37	-0.905854						
ENSECAG00000024574	GATA3	Down35	-0.907896						
ENSECAG00000017274	KRT13	Down30	-0.915744						
ENSECAG00000016829	ABTB2	Down24	-0.95072						
ENSECAG00000004091	SNN	Down20	-0.984933						
ENSECAG00000019043	SSC5D	Down19	-0.986903						
ENSECAG00000019677	EPHX1	Down17	-0.997667						
ENSECAG00000013281	IRGQ	Down16	-0.998544						
ENSECAG00000015822	VDR	Down13	-1.02021						
ENSECAG00000013201	MMP17	Down11	-1.0566						
ENSECAG00000006082	ENSECAG00000006082	Down8	-1.12489						
ENSECAG00000014911	APC2	Down7	-1.12532						
ENSECAG00000023427	HES2	Down6	-1.18056						
ENSECAG00000013662	DNAI2	Down5	-1.28415						
ENSECAG00000003025	HLX	Down2	-1.50756						
ENSECAG00000009955	TEX43			Up1	inf				
ENSECAG00000020819	AGER			Up18	1.77703				
ENSECAG00000012470	PLA2G4F			Up25	1.4647				
ENSECAG00000024712	CASR			Up34	1.27409				
ENSECAG00000020764	ENSECAG00000020764			Up42	1.07186				
ENSECAG00000019045	FGF18			Up46	1.02334				
ENSECAG00000014921	GPRC5A			Up49	1.01123				
ENSECAG00000013893	SCARA5			Up50	1.00702				
ENSECAG00000021165	STAC			Up60	0.909029				
ENSECAG00000005651	SLC35G2			Up65	0.879381				
ENSECAG00000027377	5_8S_rRNA			Up66	0.876109				
ENSECAG00000011929	NR4A3			Up76	0.819122				
ENSECAG00000016228	CA2			Up81	0.787795				
ENSECAG00000002759	MTRF1L			Up84	0.750066				
ENSECAG00000015611	ETV4			Up89	0.73455				
ENSECAG00000008811	KCNJ15			Up103	0.70262				
ENSECAG00000001815	PYURF			Up106	0.691616				
ENSECAG00000023134	KCNK5			Up140	0.592647				
ENSECAG00000011672	TMEM63C			Down59	-0.582332				
ENSECAG00000011229	SLC1A4			Down57	-0.594733				
ENSECAG00000006553	CHST4			Down52	-0.600533				
ENSECAG00000014601	GADD45A			Down51	-0.601335				
ENSECAG00000023217	KLHL17			Down50	-0.603489				
ENSECAG00000018603	SLC46A2			Down49	-0.604993				
ENSECAG00000000178	RUNX1T1			Down48	-0.609126				
ENSECAG00000023946	SDC3			Down47	-0.610372				
ENSECAG00000006847	PIK3CD			Down46	-0.613399				
ENSECAG00000000828	ENSECAG00000000828			Down43	-0.645836				
ENSECAG00000000046	FAM13C			Down40	-0.668659				
ENSECAG00000024309	LRRC24			Down39	-0.671771				
ENSECAG00000005000	PPM1K			Down38	-0.67219				
ENSECAG00000006249	ENSECAG00000006249			Down35	-0.687782				
ENSECAG00000012946	CYP7A1			Down33	-0.724282				
ENSECAG00000016753	LIMS2			Down32	-0.727509				
ENSECAG00000014767	PLXNB1			Down31	-0.727954				
ENSECAG00000008485	ITGAL			Down26	-0.766866				
ENSECAG00000020162	PRR5L			Down25	-0.79314				
ENSECAG00000023458	OLFM2			Down19	-0.876702				

APPENDIX 53: List of DEGs IN U/63-infected samples – Part 8

		U/63-WT		U/63-E186K		U/63-219		U/63-E186K-219	
Gene ID	Gene name	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC
ENSECAG00000024139	LONRF3			Down18	-0.886645				
ENSECAG00000017015	MYH11			Down17	-0.894766				
ENSECAG00000011583	KRT14			Down15	-0.945321				
ENSECAG00000011524	CAP2			Down14	-0.948603				
ENSECAG00000006058	FUT1			Down13	-0.950377				
ENSECAG00000021037	FEZF2			Down12	-0.970215				
ENSECAG00000007020	C14orf28			Down10	-0.973021				
ENSECAG00000011427	TRPM8			Down8	-1.04342				
ENSECAG00000023032	AZIN2			Down6	-1.10437				
ENSECAG00000017175	REM1			Down5	-1.11577				
ENSECAG00000012161	ZNF711			Down2	-1.14786				
ENSECAG00000012253	CXCL11					Up1	inf		
ENSECAG00000025480	SNORA71					Up2	inf		
ENSECAG00000010206	SMTNL1					Up30	1.42041		
ENSECAG00000010036	SLC15A3					Up33	1.35719		
ENSECAG00000023195	ENSECAG00000023195					Up39	1.18057		
ENSECAG00000024430	ENSECAG00000024430					Up42	1.07985		
ENSECAG00000025120	SBSN					Up44	1.05685		
ENSECAG00000006493	PTMA					Up45	1.03333		
ENSECAG00000015395	HERC5					Up51	0.944138		
ENSECAG00000019621	ENSECAG00000019621					Up53	0.936982		
ENSECAG00000008257	RBMX2					Up55	0.915059		
ENSECAG00000009273	SLC45A3					Up56	0.914838		
ENSECAG00000019590	NINL					Up57	0.906708		
ENSECAG00000010638	SREK1IP1					Up58	0.892027		
ENSECAG00000006211	TRIM21					Up59	0.876248		
ENSECAG00000022601	PARP12					Up60	0.833009		
ENSECAG00000024957	C19orf66					Up62	0.817361		
ENSECAG00000006666	MURC					Up63	0.814026		
ENSECAG00000012843	PAG1					Up65	0.776355		
ENSECAG00000018264	TRANK1					Up66	0.771467		
ENSECAG00000014880	CFAP45					Up67	0.768676		
ENSECAG00000001514	ENSECAG00000001514					Up68	0.765041		
ENSECAG00000012660	HLA-DRA					Up69	0.763677		
ENSECAG00000009586	C16orf89					Up70	0.761276		
ENSECAG00000014480	AIM2					Up73	0.725967		
ENSECAG00000004617	PML					Up74	0.719505		
ENSECAG00000006844	CALD1					Up75	0.718923		
ENSECAG00000004534	CCDC136					Up77	0.705269		
ENSECAG00000012794	LYAR					Up78	0.701695		
ENSECAG00000009383	TBCA					Up79	0.701142		
ENSECAG00000021345	POLR2K					Up80	0.684892		
ENSECAG00000002607	WBP5					Up81	0.680495		
ENSECAG00000018205	OTUD6B					Up82	0.67276		
ENSECAG00000006454	ENSECAG00000006454					Up84	0.668792		
ENSECAG00000022765	ZC3H8					Up86	0.663995		
ENSECAG00000012760	LUC7L3					Up89	0.656039		
ENSECAG00000009432	PDRG1					Up91	0.643733		
ENSECAG00000002619	GREM1					Up92	0.638078		
ENSECAG00000026989	PDCD7					Up93	0.637717		
ENSECAG00000014812	KRT23					Up94	0.636528		
ENSECAG00000002251	TLL1					Up97	0.627479		
ENSECAG00000015180	NEXN					Up99	0.620353		
ENSECAG00000011155	ATP6V1G1					Up100	0.617714		
ENSECAG00000008210	ZNFX1					Up101	0.614708		
ENSECAG00000014051	MED19					Up103	0.613226		
ENSECAG00000007848	ITGA9					Up104	0.612344		
ENSECAG00000026874	ENSECAG00000026874					Up105	0.611339		
ENSECAG00000016947	UPF3B					Up106	0.610805		
ENSECAG00000011899	PNN					Up107	0.607343		
ENSECAG00000003698	ENSECAG00000003698					Up109	0.603593		
ENSECAG00000000974	ADAR					Up110	0.602139		
ENSECAG00000009258	DAAM1					Up111	0.600076		
ENSECAG00000004232	AMER1					Up112	0.594351		
ENSECAG00000013395	RBM25					Up114	0.585833		
ENSECAG00000009310	ENSECAG00000009310					Up115	0.584806		
ENSECAG00000016025	SNRNP27					Up116	0.584254		
ENSECAG00000022972	TCHP					Up117	0.581557		
ENSECAG00000009013	CCDC107					Down98	-0.589534		
ENSECAG00000012066	FBLN5					Down97	-0.597537		
ENSECAG00000021133	FDFT1					Down96	-0.602278		
ENSECAG00000014017	PRKG1					Down95	-0.602867		
ENSECAG00000021576	SGCG					Down94	-0.603469		

APPENDIX 54: List of DEGs IN U/63-infected samples – Part 9

		U/63-WT		U/63-E186K		U/63-219		U/63-E186K-219	
Gene ID	Gene name	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC
ENSECAG00000000857	PRKCDBP					Down93	-0.609082		
ENSECAG00000004817	CDKN2C					Down92	-0.610564		
ENSECAG00000006455	CYP51A1					Down91	-0.613223		
ENSECAG00000024992	HSD17B7					Down90	-0.616537		
ENSECAG00000011479	PLAC9					Down89	-0.620718		
ENSECAG00000008926	RNF139					Down83	-0.63596		
ENSECAG00000014658	SCD					Down82	-0.638683		
ENSECAG00000024852	FABP3					Down79	-0.641854		
ENSECAG00000022769	RAB30					Down78	-0.649176		
ENSECAG00000018059	SC5D					Down76	-0.658513		
ENSECAG00000012893	ENSECAG00000012893					Down75	-0.665706		
ENSECAG00000013998	SQLE					Down74	-0.674203		
ENSECAG00000018012	SELENBP1					Down73	-0.683824		
ENSECAG00000016942	NUDT14					Down70	-0.699527		
ENSECAG00000000668	COL4A5					Down69	-0.704672		
ENSECAG00000014263	ENSECAG00000014263					Down67	-0.711492		
ENSECAG00000011271	IGFBP4					Down66	-0.717586		
ENSECAG00000026960	UBE2S					Down62	-0.735956		
ENSECAG00000005600	ZNF48					Down61	-0.750375		
ENSECAG00000004504	C19orf52					Down60	-0.756271		
ENSECAG00000009515	ATRNL1					Down59	-0.757059		
ENSECAG00000016271	CENPA					Down57	-0.761043		
ENSECAG00000023354	ENSECAG00000023354					Down55	-0.778611		
ENSECAG00000003398	MARCKS					Down53	-0.784416		
ENSECAG00000014113	CCNG2					Down51	-0.799668		
ENSECAG00000008259	ENSECAG00000008259					Down50	-0.804779		
ENSECAG00000011663	NPTX1					Down49	-0.814449		
ENSECAG00000017914	L3HYPDH					Down48	-0.8174		
ENSECAG00000015351	LRRN3					Down47	-0.82161		
ENSECAG00000023284	ENSECAG00000023284					Down46	-0.82698		
ENSECAG00000012662	SEMA3A					Down45	-0.828823		
ENSECAG00000019765	ENSECAG00000019765					Down44	-0.836411		
ENSECAG00000017407	MOXD1					Down43	-0.86484		
ENSECAG00000010554	MSMO1					Down41	-0.873197		
ENSECAG00000000795	ADHFE1					Down38	-0.894171		
ENSECAG00000022251	MAP3K15					Down36	-0.910299		
ENSECAG00000000374	PLEKHG7					Down34	-0.913772		
ENSECAG00000000154	SDHAF1					Down32	-0.919585		
ENSECAG00000021607	TXNDC8					Down31	-0.921575		
ENSECAG00000015644	SLC1A2					Down29	-0.95657		
ENSECAG00000009248	IL1RL1					Down28	-0.962654		
ENSECAG00000001140	ANAPC11					Down27	-0.967842		
ENSECAG00000018831	ENSECAG00000018831					Down26	-0.975751		
ENSECAG00000011356	ENSECAG00000011356					Down25	-0.978965		
ENSECAG00000006880	WNT16					Down22	-1.03741		
ENSECAG00000021431	ENSECAG00000021431					Down20	-1.04623		
ENSECAG00000011948	FIGF					Down19	-1.04654		
ENSECAG00000026965	SOCS1					Down18	-1.06771		
ENSECAG00000014525	ENSECAG00000014525					Down17	-1.09085		
ENSECAG00000021076	MBNL3					Down16	-1.10087		
ENSECAG00000015692	SERTAD4					Down14	-1.16276		
ENSECAG00000024192	FAM180A					Down12	-1.22548		
ENSECAG00000011313	SLC35F1					Down10	-1.24767		
ENSECAG00000007071	ENSECAG00000007071					Down9	-1.28301		
ENSECAG00000011090	PTPN18					Down7	-1.33035		
ENSECAG00000006174	ENSECAG00000006174					Down6	-1.33559		
ENSECAG00000018591	CTU1					Down5	-1.56674		
ENSECAG00000013137	ENSECAG00000013137					Down4	-1.68257		
ENSECAG00000027702	ENSECAG00000027702					Down1	-inf		
ENSECAG00000021420	PSORS1C2							Up1	inf
ENSECAG00000027406	SNORA31							Up3	inf
ENSECAG00000018077	ENSECAG00000018077							Up20	1.78161
ENSECAG00000003848	ENSECAG00000003848							Up22	1.74835
ENSECAG00000015232	ZNF619							Up23	1.61786
ENSECAG00000001216	ENSECAG00000001216							Up24	1.60188
ENSECAG00000023287	ENSECAG00000023287							Up30	1.51745
ENSECAG00000015950	CCDC68							Up32	1.46019
ENSECAG00000003073	ENSECAG00000003073							Up36	1.36806
ENSECAG00000020282	FAM71F2							Up37	1.36036
ENSECAG00000020088	NOSTRIN							Up38	1.35307
ENSECAG00000012949	ENSECAG00000012949							Up39	1.34644
ENSECAG00000020053	SLC4A4							Up40	1.32211
ENSECAG00000010739	ITGA2							Up41	1.31901

APPENDIX 55: List of DEGs IN U/63-infected samples – Part 10

		U/63-WT		U/63-E186K		U/63-219		U/63-E186K-219	
Gene ID	Gene name	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC
ENSECAG00000005115	ZNF454							Up42	1.31614
ENSECAG00000006323	LEKR1							Up46	1.29041
ENSECAG00000010696	ENSECAG00000010696							Up48	1.27549
ENSECAG00000016295	EED							Up49	1.27278
ENSECAG00000013869	CFI							Up51	1.26099
ENSECAG00000021404	ENSECAG00000021404							Up56	1.21353
ENSECAG00000014092	ENSECAG00000014092							Up59	1.19261
ENSECAG00000007757	DNAJC15							Up60	1.16908
ENSECAG00000003510	TXNDC9							Up61	1.16618
ENSECAG00000020127	THEM4							Up62	1.16257
ENSECAG00000015237	PIBF1							Up63	1.15827
ENSECAG00000005951	TMEM81							Up66	1.13633
ENSECAG00000016771	ENSECAG00000016771							Up67	1.1355
ENSECAG00000021582	ENSECAG00000021582							Up70	1.12995
ENSECAG00000017181	PTGS2							Up73	1.12235
ENSECAG00000012316	DTWD1							Up76	1.09488
ENSECAG00000021942	TEX30							Up77	1.09118
ENSECAG00000013268	AP1S3							Up78	1.08918
ENSECAG00000014388	ZNF432							Up80	1.08198
ENSECAG00000001069	CCRL2							Up81	1.07963
ENSECAG00000001696	ENSECAG00000001696							Up82	1.07929
ENSECAG00000011685	ENSECAG00000011685							Up83	1.06848
ENSECAG00000024704	CETN3							Up85	1.06309
ENSECAG00000023859	BTG3							Up86	1.06205
ENSECAG00000000785	ENSECAG00000000785							Up87	1.06037
ENSECAG00000000546	CCDC62							Up88	1.06035
ENSECAG00000018128	ENSECAG00000018128							Up91	1.05572
ENSECAG00000017138	ENSECAG00000017138							Up95	1.04454
ENSECAG00000018881	ENSECAG00000018881							Up98	1.03269
ENSECAG00000023471	ZNF879							Up100	1.02874
ENSECAG00000009776	CEP57L1							Up101	1.02754
ENSECAG00000008540	C5orf28							Up103	1.01649
ENSECAG00000000629	ENSECAG00000000629							Up104	1.01507
ENSECAG00000026852	ZCCHC10							Up107	1.00462
ENSECAG00000019742	CCDC126							Up109	1.00252
ENSECAG00000010663	SKIL							Up110	0.999293
ENSECAG00000002936	WDR89							Up112	0.997324
ENSECAG00000014054	CDKL4							Up117	0.990473
ENSECAG00000014144	ENSECAG00000014144							Up118	0.989901
ENSECAG00000018495	SSB							Up119	0.98935
ENSECAG00000019188	ESF1							Up121	0.987375
ENSECAG00000010084	DNAJC10							Up122	0.986149
ENSECAG00000016480	SPCS3							Up123	0.985784
ENSECAG00000000750	MMP3							Up126	0.981314
ENSECAG00000000473	ATAD1							Up127	0.980964
ENSECAG00000023644	PPAT							Up128	0.978733
ENSECAG00000019554	LRRC40							Up130	0.97854
ENSECAG00000017606	SNAPC1							Up131	0.978287
ENSECAG00000012601	CEP85L							Up132	0.974621
ENSECAG00000010744	ADSS							Up133	0.974107
ENSECAG00000020144	UBXN2A							Up134	0.972839
ENSECAG00000010608	ENSECAG00000010608							Up137	0.968727
ENSECAG00000006518	SLC39A10							Up138	0.96839
ENSECAG00000005272	ENSECAG00000005272							Up139	0.966818
ENSECAG00000011946	CHEK1							Up141	0.965269
ENSECAG00000013569	LAMTOR3							Up143	0.961549
ENSECAG00000011218	HAT1							Up144	0.960955
ENSECAG00000021075	FGD6							Up147	0.955674
ENSECAG00000024815	RFESD							Up148	0.955448
ENSECAG00000018428	HSPH1							Up150	0.954381
ENSECAG00000022011	AKAP7							Up151	0.954176
ENSECAG00000000733	C18orf21							Up154	0.947336
ENSECAG00000000794	EMB							Up156	0.936423
ENSECAG00000015765	THAP6							Up157	0.935283
ENSECAG00000009502	AASDH							Up159	0.933024
ENSECAG00000024018	TWF1							Up160	0.92771
ENSECAG00000004023	SSBP1							Up161	0.926606
ENSECAG00000026810	FAM133B							Up162	0.926265
ENSECAG00000014393	CEP135							Up165	0.923896
ENSECAG00000014702	MEDAG							Up166	0.923659
ENSECAG00000009859	PAPOLG							Up167	0.920183
ENSECAG00000014665	ASF1A							Up168	0.915956
ENSECAG00000011381	TRIM13							Up169	0.915786

APPENDIX 56: List of DEGs IN U/63-infected samples – Part 11

		U/63-WT		U/63-E186K		U/63-219		U/63-E186K-219	
Gene ID	Gene name	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC
ENSECAG00000014890	BRCA2							Up170	0.914551
ENSECAG00000011465	ISCA1							Up173	0.908385
ENSECAG00000011151	USP1							Up174	0.907819
ENSECAG00000013901	ANKRD12							Up175	0.907316
ENSECAG00000011563	GXYLT1							Up176	0.906656
ENSECAG00000021329	CHRNA5							Up177	0.905666
ENSECAG00000008226	NPAT							Up178	0.905321
ENSECAG00000018407	SNRPB2							Up180	0.90142
ENSECAG00000002222	ZBTB6							Up181	0.898987
ENSECAG00000017788	TRMT11							Up183	0.894171
ENSECAG00000000971	SUCO							Up184	0.893548
ENSECAG00000013226	ZDHHC6							Up185	0.892831
ENSECAG00000021782	DCK							Up186	0.892166
ENSECAG00000008096	FAM111B							Up187	0.890394
ENSECAG00000022016	NAPB							Up189	0.889208
ENSECAG00000009554	SLC25A40							Up190	0.886986
ENSECAG00000020387	MRPS18C							Up191	0.886957
ENSECAG00000008461	DEPDC1							Up192	0.886739
ENSECAG00000024416	ANAPC10							Up193	0.885864
ENSECAG00000016106	AGPAT5							Up194	0.885861
ENSECAG00000022164	SMC4							Up195	0.884991
ENSECAG00000010606	CCDC169							Up196	0.884799
ENSECAG00000016676	SEMA6D							Up197	0.884087
ENSECAG00000018454	RBBP8							Up199	0.883905
ENSECAG00000009303	GFPT2							Up200	0.883629
ENSECAG00000008428	CEP83							Up202	0.882909
ENSECAG00000010845	RASA2							Up203	0.881371
ENSECAG00000004370	ENSECAG00000004370							Up204	0.87915
ENSECAG00000008716	EIF3J							Up205	0.878892
ENSECAG00000018870	IVNS1ABP							Up206	0.877243
ENSECAG00000021185	ZFP82							Up208	0.87433
ENSECAG00000018478	NIFK							Up209	0.873967
ENSECAG00000000609	BIRC2							Up210	0.873165
ENSECAG00000024267	TFRC							Up212	0.869865
ENSECAG00000019388	MPHOSPH6							Up214	0.868697
ENSECAG00000026943	ZFAND1							Up215	0.868481
ENSECAG00000004919	CACYBP							Up216	0.867527
ENSECAG00000008261	FCF1							Up219	0.864592
ENSECAG00000010746	RAD18							Up220	0.864507
ENSECAG00000018267	CENPC							Up221	0.864364
ENSECAG00000020337	DOCK11							Up222	0.86312
ENSECAG00000011889	USPL1							Up226	0.857688
ENSECAG00000019146	NDUFAF6							Up227	0.856812
ENSECAG00000023133	ENSECAG00000023133							Up229	0.856381
ENSECAG00000021440	TBC1D8B							Up230	0.853147
ENSECAG00000006038	SGOL2							Up231	0.852325
ENSECAG00000016705	UGP2							Up234	0.851063
ENSECAG00000022387	RUNDC3B							Up235	0.850115
ENSECAG00000022541	ZNF25							Up236	0.849159
ENSECAG00000024999	SMC5							Up238	0.848589
ENSECAG00000013725	NEDD1							Up240	0.84797
ENSECAG00000016586	NUCB2							Up241	0.8477
ENSECAG00000021200	ENSECAG00000021200							Up242	0.846875
ENSECAG00000007436	FAM210A							Up243	0.845835
ENSECAG00000010070	ZNF484							Up244	0.844603
ENSECAG00000024762	RECQL							Up246	0.843354
ENSECAG00000009447	KIAA1586							Up247	0.843326
ENSECAG00000017190	ENSECAG00000017190							Up248	0.843295
ENSECAG00000016109	RCHY1							Up249	0.842404
ENSECAG00000023808	ATAD2							Up251	0.842176
ENSECAG00000013631	SPOPL							Up252	0.841639
ENSECAG00000013523	ZUFSP							Up253	0.841258
ENSECAG00000005234	ENSECAG00000005234							Up254	0.840729
ENSECAG00000024810	PLA2G4A							Up255	0.840594
ENSECAG00000007411	RPAP2							Up256	0.839565
ENSECAG00000009638	MPP6							Up257	0.839056
ENSECAG00000019829	FNTA							Up258	0.838624
ENSECAG00000021949	RAD54B							Up260	0.837896
ENSECAG00000016351	ENSECAG00000016351							Up261	0.837601
ENSECAG00000019865	ZNF277							Up262	0.836582
ENSECAG00000016479	ENSECAG00000016479							Up263	0.835472
ENSECAG00000015597	HIF1A							Up264	0.833704
ENSECAG00000003122	FRG1							Up265	0.833678

APPENDIX 57: List of DEGs IN U/63-infected samples – Part 12

		U/63-WT		U/63-E186K		U/63-219		U/63-E186K-219	
Gene ID	Gene name	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC
ENSECAG00000018650	SMC2							Up266	0.832745
ENSECAG00000017007	FANCL							Up269	0.831852
ENSECAG00000027684	ND1							Up270	0.831689
ENSECAG00000008390	FOXN2							Up273	0.830815
ENSECAG00000012475	SGOL1							Up275	0.830233
ENSECAG00000015468	MGAT4A							Up276	0.830121
ENSECAG00000012780	CBX3							Up278	0.828811
ENSECAG00000015297	PRIM1							Up280	0.82814
ENSECAG00000013255	ENSECAG00000013255							Up281	0.826572
ENSECAG00000023208	NAA50							Up282	0.82613
ENSECAG00000013326	POC1B							Up283	0.825179
ENSECAG00000021679	CDK17							Up284	0.824436
ENSECAG00000011928	LRRC8C							Up285	0.824414
ENSECAG00000023941	PIK3CB							Up286	0.823712
ENSECAG00000005550	SFN							Up287	0.822724
ENSECAG00000024396	LMBR1							Up289	0.821743
ENSECAG00000019455	NIPAL1							Up290	0.820077
ENSECAG00000007814	ENSECAG00000007814							Up291	0.819874
ENSECAG00000014613	DBF4							Up293	0.81887
ENSECAG00000018792	ZNF782							Up295	0.817939
ENSECAG00000024709	COPS2							Up297	0.816191
ENSECAG00000005178	CCSAP							Up298	0.816035
ENSECAG00000018274	FAM172A							Up299	0.815964
ENSECAG00000005384	SPRY4							Up300	0.813706
ENSECAG00000012685	SLC10A6							Up301	0.813273
ENSECAG00000010304	ZNF644							Up302	0.813061
ENSECAG00000024588	ENSECAG00000024588							Up303	0.8124
ENSECAG00000008892	ENSECAG00000008892							Up304	0.812272
ENSECAG00000015330	SRFBP1							Up305	0.811695
ENSECAG00000022748	ZNF850							Up306	0.811598
ENSECAG00000019890	VRK1							Up307	0.811203
ENSECAG00000015993	ENSECAG00000015993							Up308	0.81055
ENSECAG00000021912	C5orf34							Up309	0.81055
ENSECAG00000021009	GUCY1B3							Up311	0.809782
ENSECAG00000013562	ZNF706							Up312	0.809778
ENSECAG00000016028	N4BP2L2							Up313	0.808837
ENSECAG00000021480	ZNF573							Up315	0.808623
ENSECAG00000017104	ERRFI1							Up316	0.808531
ENSECAG00000006526	CENPJ							Up317	0.807766
ENSECAG00000010929	MRPS35							Up318	0.807129
ENSECAG00000014888	NEK5							Up319	0.806471
ENSECAG00000009568	CENPI							Up320	0.806382
ENSECAG00000003722	ZFP62							Up321	0.806048
ENSECAG00000022278	CASC5							Up322	0.805769
ENSECAG00000008456	SLC39A6							Up323	0.805328
ENSECAG00000017041	ACTR6							Up324	0.804921
ENSECAG00000022169	PDSS1							Up325	0.804535
ENSECAG00000007552	HYKK							Up326	0.804533
ENSECAG00000020404	DENND1B							Up328	0.803833
ENSECAG00000021030	IFT74							Up329	0.802744
ENSECAG00000013653	TPMT							Up330	0.802692
ENSECAG00000016302	ENSECAG00000016302							Up331	0.799943
ENSECAG00000018635	ZNF674							Up332	0.799163
ENSECAG00000000543	CYBRD1							Up333	0.798916
ENSECAG00000023709	SRSF10							Up334	0.795492
ENSECAG00000012133	VEGFC							Up335	0.795275
ENSECAG00000018299	FKBP3							Up336	0.795122
ENSECAG00000000194	ZMYM5							Up337	0.794542
ENSECAG00000020631	CISD2							Up338	0.794491
ENSECAG00000027675	ND4							Up339	0.79389
ENSECAG00000016313	PTPN4							Up340	0.793514
ENSECAG00000017951	GTF2E2							Up341	0.793424
ENSECAG00000015910	KTN1							Up342	0.791806
ENSECAG00000017121	RPF2							Up343	0.791251
ENSECAG00000009654	DCAF13							Up344	0.789531
ENSECAG00000020873	GIN1							Up345	0.789478
ENSECAG00000010922	RAD50							Up346	0.789154
ENSECAG00000013909	SIRT1							Up347	0.788563
ENSECAG00000013041	ETF1							Up348	0.787588
ENSECAG00000007351	PLAU							Up349	0.787434
ENSECAG00000013757	ERO1A							Up350	0.787069
ENSECAG00000016530	LYPLA1							Up351	0.78654
ENSECAG00000013770	INTS6							Up352	0.786222

APPENDIX 58: List of DEGs IN U/63-infected samples – Part 13

		U/63-WT		U/63-E186K		U/63-219		U/63-E186K-219	
Gene ID	Gene name	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC
ENSECAG00000021776	ENSECAG00000021776							Up353	0.786217
ENSECAG00000013889	NBN							Up354	0.786073
ENSECAG00000005243	MOCS2							Up355	0.785907
ENSECAG00000008211	ENSECAG00000008211							Up356	0.784995
ENSECAG00000014325	LIN9							Up357	0.784569
ENSECAG00000023994	DIS3							Up358	0.784169
ENSECAG00000008568	SLC35A3							Up359	0.783492
ENSECAG00000001407	CHSY3							Up360	0.782537
ENSECAG00000020958	ARL6							Up361	0.78164
ENSECAG00000019325	CNOT6							Up362	0.780968
ENSECAG00000019228	UBLCP1							Up363	0.780914
ENSECAG00000024277	HLTF							Up364	0.780794
ENSECAG00000008642	PRKD3							Up365	0.780113
ENSECAG00000014173	ZRANB2							Up366	0.778485
ENSECAG00000018207	RMND1							Up368	0.777712
ENSECAG00000012085	SWT1							Up369	0.777512
ENSECAG00000023113	SASS6							Up370	0.776898
ENSECAG00000000767	SLC4A7							Up371	0.776671
ENSECAG00000000823	ATG4C							Up374	0.776204
ENSECAG00000023747	CCBL2							Up375	0.776063
ENSECAG00000000188	NMD3							Up377	0.77561
ENSECAG00000018129	PSMC6							Up378	0.77556
ENSECAG00000004339	ENSECAG00000004339							Up379	0.774834
ENSECAG00000027699	ENSECAG00000027699							Up380	0.774814
ENSECAG00000018042	SRP54							Up381	0.774776
ENSECAG00000008289	C1orf27							Up382	0.774672
ENSECAG00000024373	ZBTB11							Up383	0.774646
ENSECAG00000005653	PTP4A2							Up384	0.774479
ENSECAG00000010986	MTDH							Up385	0.774228
ENSECAG00000014894	FASTKD3							Up386	0.772111
ENSECAG00000013441	APOO							Up387	0.772022
ENSECAG00000013223	PBK							Up388	0.77197
ENSECAG00000009163	NUF2							Up389	0.771462
ENSECAG00000021732	GOSR1							Up390	0.771193
ENSECAG00000006638	CASP8AP2							Up391	0.771134
ENSECAG00000019097	TANK							Up392	0.770936
ENSECAG00000013090	POLA1							Up393	0.770924
ENSECAG00000009697	ENSECAG00000009697							Up394	0.770849
ENSECAG00000013053	TOP2A							Up395	0.770577
ENSECAG00000014475	STK39							Up396	0.770064
ENSECAG00000019782	PNISR							Up397	0.77001
ENSECAG00000017978	CCNYL1							Up398	0.769857
ENSECAG00000008222	ENSECAG00000008222							Up399	0.769832
ENSECAG00000015355	CLK1							Up400	0.769787
ENSECAG00000013845	HUS1							Up401	0.769061
ENSECAG00000020240	ARPP19							Up402	0.768674
ENSECAG00000011527	CASD1							Up403	0.76783
ENSECAG00000024764	ARHGAP18							Up404	0.767817
ENSECAG00000017193	PUM3							Up405	0.767812
ENSECAG00000010800	PPID							Up406	0.767061
ENSECAG00000010736	KIAA1551							Up408	0.764425
ENSECAG00000019984	ZCWPW2							Up409	0.763664
ENSECAG00000002021	ENSECAG00000002021							Up410	0.763293
ENSECAG00000022265	SMC6							Up411	0.763122
ENSECAG00000009632	PSMA4							Up413	0.762404
ENSECAG00000011545	RNF149							Up414	0.762032
ENSECAG00000014668	ZNF614							Up415	0.761629
ENSECAG00000016651	NSMCE4A							Up416	0.761559
ENSECAG00000013040	MANEA							Up417	0.761007
ENSECAG00000012903	WDR76							Up418	0.760876
ENSECAG00000009760	GMNN							Up419	0.759888
ENSECAG00000008106	FAM111A							Up420	0.759822
ENSECAG00000008396	CALM2							Up421	0.759687
ENSECAG00000013138	RAB27B							Up422	0.759409
ENSECAG00000008088	NIPSNAP3B							Up423	0.759158
ENSECAG00000013691	MALT1							Up424	0.758805
ENSECAG00000018335	ERCC8							Up425	0.758771
ENSECAG00000000237	DGKH							Up426	0.758484
ENSECAG00000022599	RFC4							Up427	0.758483
ENSECAG00000024978	ZNHIT6							Up428	0.758462
ENSECAG00000005765	ADI1							Up429	0.758343
ENSECAG00000023790	MNAT1							Up430	0.758137
ENSECAG00000020788	MECOM							Up431	0.757715

APPENDIX 59: List of DEGs IN U/63-infected samples – Part 14

		U/63-WT		U/63-E186K		U/63-219		U/63-E186K-219	
Gene ID	Gene name	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC
ENSECAG00000022617	PRKAA1							Up432	0.75737
ENSECAG00000011592	HNMT							Up433	0.756717
ENSECAG00000011108	PARP8							Up434	0.756694
ENSECAG00000019257	RBPJ							Up435	0.756265
ENSECAG00000013766	E2F5							Up436	0.756186
ENSECAG00000007770	ERCC6L2							Up438	0.755168
ENSECAG00000000558	KIF5B							Up439	0.754153
ENSECAG00000015412	WDHD1							Up441	0.753124
ENSECAG00000015098	MORC3							Up444	0.751967
ENSECAG00000002866	UGCG							Up445	0.751774
ENSECAG00000004756	METTL5							Up446	0.751608
ENSECAG00000018130	ZFAND6							Up447	0.751001
ENSECAG00000009676	SAP30							Up448	0.750959
ENSECAG00000010978	ENSECAG00000010978							Up449	0.75092
ENSECAG00000012777	MPHOSPH8							Up450	0.750868
ENSECAG00000022460	U2SURP							Up451	0.750583
ENSECAG00000018885	ENSECAG00000018885							Up452	0.750449
ENSECAG00000014852	DUS4L							Up454	0.75017
ENSECAG00000009898	UEVLD							Up455	0.750152
ENSECAG00000000308	RICTOR							Up456	0.749504
ENSECAG00000010191	PYGO1							Up458	0.749213
ENSECAG00000010482	CCDC186							Up459	0.748646
ENSECAG00000000648	LCORL							Up460	0.748418
ENSECAG00000011878	DTL							Up461	0.747845
ENSECAG00000024976	OCIAD1							Up462	0.746897
ENSECAG00000010433	CHMP2B							Up463	0.746774
ENSECAG00000022396	SMC3							Up464	0.746237
ENSECAG00000010054	NOX4							Up465	0.745877
ENSECAG00000001533	TRMT10C							Up466	0.744993
ENSECAG00000024558	ENSECAG00000024558							Up467	0.744984
ENSECAG00000023959	GTF2H3							Up468	0.74478
ENSECAG00000022582	FBXW7							Up469	0.744552
ENSECAG00000021058	RMDN2							Up470	0.743931
ENSECAG00000011139	ENSECAG00000011139							Up471	0.743769
ENSECAG00000013324	TXN							Up472	0.743567
ENSECAG00000022045	DNTTIP2							Up474	0.73877
ENSECAG00000023445	NUDCD1							Up475	0.738599
ENSECAG00000016383	USP53							Up476	0.736881
ENSECAG00000020276	HOOK3							Up478	0.735647
ENSECAG00000024279	FAM200B							Up479	0.735024
ENSECAG00000004389	MFF							Up480	0.734382
ENSECAG00000024495	LYSMD3							Up481	0.734222
ENSECAG00000002556	ZNF770							Up482	0.734039
ENSECAG00000027676	ENSECAG00000027676							Up483	0.733823
ENSECAG00000002402	ENSECAG00000002402							Up484	0.733424
ENSECAG00000017714	FDX1							Up485	0.733198
ENSECAG00000017271	ARHGAP5							Up486	0.732418
ENSECAG00000022168	ENSECAG00000022168							Up487	0.732399
ENSECAG00000017588	MMD							Up488	0.732394
ENSECAG00000024964	ATP6V1C1							Up489	0.732385
ENSECAG00000015920	PHTF2							Up490	0.732208
ENSECAG00000019398	TFPI2							Up491	0.730831
ENSECAG00000020674	NDUFB3							Up492	0.730739
ENSECAG00000013644	TRDMT1							Up495	0.729582
ENSECAG00000004651	RAP1B							Up496	0.728989
ENSECAG00000006779	USP12							Up497	0.728922
ENSECAG00000015560	MOB4							Up498	0.728632
ENSECAG00000018064	CCDC82							Up499	0.728617
ENSECAG00000011935	SMURF2							Up500	0.727753
ENSECAG00000019335	ENSECAG00000019335							Up501	0.727541
ENSECAG00000000306	GLRX3							Up502	0.727311
ENSECAG00000002063	C2orf76							Up503	0.727143
ENSECAG00000021355	STAM2							Up504	0.726268
ENSECAG00000015019	PIK3CA							Up505	0.725859
ENSECAG00000016922	AZIN1							Up506	0.725054
ENSECAG00000014227	ENSECAG00000014227							Up507	0.724213
ENSECAG00000002691	RSBN1L							Up508	0.723845
ENSECAG00000014420	TRNT1							Up509	0.723566
ENSECAG00000020878	ENSECAG00000020878							Up510	0.723345
ENSECAG00000024392	BRIP1							Up511	0.72263
ENSECAG00000011295	THAP2							Up512	0.722421
ENSECAG00000002452	CCDC121							Up513	0.722386
ENSECAG00000002675	ENSECAG00000002675							Up514	0.722304

APPENDIX 60: List of DEGs IN U/63-infected samples – Part 15

		U/63-WT		U/63-E186K		U/63-219		U/63-E186K-219	
Gene ID	Gene name	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC
ENSECAG00000010367	RAD51AP1							Up515	0.722242
ENSECAG00000021408	VRK2							Up516	0.722188
ENSECAG00000018694	FANCM							Up517	0.722029
ENSECAG00000018114	ENSECAG00000018114							Up518	0.720715
ENSECAG00000023331	KIF20B							Up519	0.720676
ENSECAG00000007966	NCK1							Up520	0.720543
ENSECAG00000015497	NAA16							Up521	0.720495
ENSECAG00000016974	BRCA1							Up522	0.719901
ENSECAG00000006911	ENSECAG00000006911							Up523	0.719855
ENSECAG00000009794	ENSECAG00000009794							Up524	0.719602
ENSECAG00000006497	SETD9							Up525	0.719387
ENSECAG00000024180	BTF3L4							Up526	0.718872
ENSECAG00000024327	FBXL5							Up527	0.718589
ENSECAG00000019841	CYB5R4							Up528	0.718263
ENSECAG00000014112	SLC20A1							Up529	0.717378
ENSECAG00000017169	ZNF283							Up531	0.717014
ENSECAG00000019810	KIF2A							Up532	0.716804
ENSECAG00000017911	PDCD10							Up533	0.716619
ENSECAG00000019791	RAB11FIP2							Up534	0.716448
ENSECAG00000014686	HAUS6							Up535	0.715297
ENSECAG00000007207	SLC35A1							Up536	0.715123
ENSECAG00000021538	CKAP2L							Up537	0.715088
ENSECAG00000023723	SKIV2L2							Up539	0.714449
ENSECAG00000010564	GLMN							Up540	0.714396
ENSECAG00000003883	CDC26							Up541	0.714359
ENSECAG00000013285	ATR							Up542	0.714123
ENSECAG00000008693	PTPN12							Up543	0.714063
ENSECAG00000022920	VPS54							Up544	0.713625
ENSECAG00000013488	TMEM38B							Up545	0.713571
ENSECAG00000012634	CEP78							Up546	0.713163
ENSECAG00000022376	ENSECAG00000022376							Up547	0.711951
ENSECAG00000014937	ENSECAG00000014937							Up548	0.710822
ENSECAG00000015796	LIAS							Up549	0.710226
ENSECAG00000013436	TCEA1							Up550	0.709893
ENSECAG00000022494	USP15							Up552	0.708881
ENSECAG00000016556	LRP12							Up553	0.708484
ENSECAG00000022698	CEP162							Up554	0.708482
ENSECAG00000017727	ORC3							Up555	0.708037
ENSECAG00000011322	RPS6KA3							Up556	0.707851
ENSECAG00000005798	CNOT7							Up557	0.707494
ENSECAG00000018523	RNMT							Up558	0.707067
ENSECAG00000017971	DDX43							Up559	0.706771
ENSECAG00000008524	SSBIP1							Up560	0.706739
ENSECAG00000001545	TSPAN12							Up562	0.706078
ENSECAG00000000477	ENSECAG00000000477							Up564	0.705457
ENSECAG00000021154	ZNF699							Up565	0.705208
ENSECAG00000024383	RARS							Up566	0.705148
ENSECAG00000026906	COX16							Up568	0.705042
ENSECAG00000019205	PPP2R2A							Up569	0.704791
ENSECAG00000004180	HSPA6							Up570	0.704738
ENSECAG00000018438	CEP55							Up571	0.703252
ENSECAG00000018686	PPA1							Up572	0.703212
ENSECAG00000009967	PHAX							Up573	0.702434
ENSECAG00000021700	FLT1							Up574	0.70149
ENSECAG00000012624	POLD3							Up575	0.701158
ENSECAG00000023263	RNF13							Up576	0.701055
ENSECAG00000022108	BORA							Up577	0.700225
ENSECAG00000017215	ZFP30							Up578	0.699635
ENSECAG00000024648	EPC1							Up579	0.699336
ENSECAG00000004011	ENSECAG00000004011							Up580	0.698884
ENSECAG00000023186	ZCCHC9							Up581	0.697818
ENSECAG00000026853	CEP152							Up582	0.696754
ENSECAG00000000102	EIF5B							Up583	0.696744
ENSECAG00000016168	CLIC4							Up584	0.696715
ENSECAG00000023076	GPCPD1							Up586	0.696244
ENSECAG00000022536	REST							Up588	0.695712
ENSECAG00000023697	FNDC3A							Up589	0.695461
ENSECAG00000001496	ZNF407							Up590	0.695444
ENSECAG00000020412	LDAH							Up591	0.693903
ENSECAG00000014326	SNX25							Up592	0.693711
ENSECAG00000024525	ENSECAG00000024525							Up593	0.693372
ENSECAG00000009283	ZNF114							Up594	0.693359
ENSECAG00000017115	ENSECAG00000017115							Up595	0.693052

APPENDIX 61: List of DEGs IN U/63-infected samples – Part 16

		U/63-WT		U/63-E186K		U/63-219		U/63-E186K-219	
Gene ID	Gene name	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC
ENSECAG00000021693	CAPZA2							Up596	0.692931
ENSECAG00000013427	RAB23							Up597	0.692732
ENSECAG00000015025	GCH1							Up598	0.692339
ENSECAG00000017850	TAF1D							Up599	0.691947
ENSECAG00000027671	ATP6							Up600	0.691656
ENSECAG00000015326	RFC5							Up601	0.691608
ENSECAG00000024811	TSC22D2							Up602	0.691029
ENSECAG00000023244	TAF1B							Up603	0.6906
ENSECAG00000009655	FAM60A							Up604	0.690235
ENSECAG00000009119	JAK2							Up605	0.690057
ENSECAG00000012157	YAF2							Up606	0.689951
ENSECAG00000015589	MTM1							Up608	0.689677
ENSECAG00000022933	ELL2							Up607	0.689677
ENSECAG00000016494	MICU2							Up609	0.68907
ENSECAG00000007680	RPA3							Up610	0.688894
ENSECAG00000026939	GINS1							Up611	0.688713
ENSECAG00000016387	NDC1							Up612	0.6886
ENSECAG00000000090	MAP3K8							Up613	0.68833
ENSECAG00000021027	PAPOLA							Up614	0.688188
ENSECAG00000012409	SAR1B							Up615	0.687703
ENSECAG00000000426	CENPL							Up616	0.687685
ENSECAG00000012484	NAMPT							Up617	0.687526
ENSECAG00000011326	TAPT1							Up618	0.687366
ENSECAG00000014323	SKA3							Up619	0.687278
ENSECAG00000019428	UTP15							Up620	0.685909
ENSECAG00000017368	MTF2							Up621	0.6856
ENSECAG00000021001	SENP6							Up622	0.685515
ENSECAG00000015041	CEP76							Up623	0.685249
ENSECAG00000021033	WBP4							Up624	0.684971
ENSECAG00000022573	USP25							Up625	0.684307
ENSECAG00000014865	ADGRA3							Up626	0.684286
ENSECAG00000023648	ABCE1							Up627	0.684255
ENSECAG00000006626	ZDHHC17							Up628	0.684243
ENSECAG00000006103	NIPSNAP3A							Up629	0.6838
ENSECAG00000009108	C1orf189							Up630	0.683627
ENSECAG00000012597	EIF5							Up631	0.683592
ENSECAG00000000110	WRN							Up632	0.683582
ENSECAG00000008999	CHORDC1							Up633	0.68321
ENSECAG00000016131	FRMD4B							Up634	0.682452
ENSECAG00000017704	ZBTB1							Up635	0.681805
ENSECAG00000009162	ENSECAG00000009162							Up636	0.680927
ENSECAG00000015739	TOPBP1							Up637	0.680471
ENSECAG00000012168	RCOR1							Up638	0.67923
ENSECAG00000020566	DYNC2LI1							Up639	0.679029
ENSECAG00000000889	ZNF322							Up640	0.678948
ENSECAG00000019842	NAA15							Up641	0.67894
ENSECAG00000023625	FCHO2							Up642	0.678939
ENSECAG00000020208	LIFR							Up643	0.67888
ENSECAG00000005678	TMEM200A							Up644	0.678736
ENSECAG00000021484	TMEM2							Up645	0.678515
ENSECAG00000021834	MIOS							Up646	0.678383
ENSECAG00000023851	SCAF11							Up647	0.678322
ENSECAG00000017348	BCAP29							Up648	0.678143
ENSECAG00000023456	NDUFAF4							Up649	0.677993
ENSECAG00000014530	SEPT10							Up650	0.67784
ENSECAG00000008584	AKAP10							Up651	0.677457
ENSECAG00000016850	KRIT1							Up652	0.677206
ENSECAG00000013476	PHF6							Up653	0.676845
ENSECAG00000009507	ENSECAG00000009507							Up654	0.676558
ENSECAG00000002569	FLRT3							Up655	0.676513
ENSECAG00000008493	SERPINB10							Up656	0.676204
ENSECAG00000026976	ENSECAG00000026976							Up657	0.67612
ENSECAG00000009872	ROCK2							Up658	0.675653
ENSECAG00000007487	ENSECAG00000007487							Up659	0.675001
ENSECAG00000011706	DENND6A							Up660	0.67485
ENSECAG00000013772	MDFIC							Up661	0.674618
ENSECAG00000014496	PPP2R3A							Up662	0.674427
ENSECAG00000024369	NIP7							Up663	0.674205
ENSECAG00000020392	AZI2							Up664	0.674063
ENSECAG00000015902	ZNF615							Up665	0.673927
ENSECAG00000014841	ENSECAG00000014841							Up666	0.673717
ENSECAG00000008458	ENSECAG00000008458							Up667	0.673458
ENSECAG00000010560	CXCL2							Up668	0.673179

APPENDIX 62: List of DEGs IN U/63-infected samples – Part 17

		U/63-WT		U/63-E186K		U/63-219		U/63-E186K-219	
Gene ID	Gene name	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC
ENSECAG00000009214	SACM1L							Up669	0.672708
ENSECAG00000024696	NOP58							Up670	0.671405
ENSECAG00000005916	CEP44							Up671	0.671225
ENSECAG00000004043	SOCS4							Up672	0.671163
ENSECAG00000000757	ENSECAG00000000757							Up673	0.671138
ENSECAG00000015185	AK6							Up674	0.670738
ENSECAG00000011712	C5orf22							Up675	0.670561
ENSECAG00000015373	ENSECAG00000015373							Up676	0.670422
ENSECAG00000017560	USP33							Up677	0.670337
ENSECAG00000006749	ATP13A3							Up678	0.670198
ENSECAG00000015452	MCM8							Up679	0.670004
ENSECAG00000017376	NEMF							Up680	0.669359
ENSECAG00000006890	MAP3K1							Up681	0.668971
ENSECAG00000017534	HPRT1							Up682	0.668674
ENSECAG00000024909	FBXL2							Up683	0.668606
ENSECAG00000019135	SRSF12							Up684	0.668479
ENSECAG00000008011	MLH3							Up685	0.668224
ENSECAG00000020148	SGCB							Up686	0.667324
ENSECAG00000002798	KIFAP3							Up687	0.667251
ENSECAG00000022066	SCYL2							Up689	0.666742
ENSECAG00000007007	STT3B							Up690	0.666268
ENSECAG00000015352	PPP2R3C							Up691	0.666247
ENSECAG00000019624	OSTF1							Up692	0.666028
ENSECAG00000024244	CASP10							Up693	0.665918
ENSECAG00000016769	TAF4B							Up694	0.665667
ENSECAG00000000313	ECI2							Up695	0.665365
ENSECAG00000018413	NUP43							Up696	0.664509
ENSECAG00000018209	CAAP1							Up697	0.664173
ENSECAG00000018367	ZWILCH							Up698	0.663954
ENSECAG00000014764	FAM206A							Up699	0.663847
ENSECAG00000004124	MMGT1							Up700	0.663807
ENSECAG00000006587	RWDD1							Up701	0.663734
ENSECAG00000011198	ENSECAG00000011198							Up702	0.663689
ENSECAG00000014344	ZNF568							Up703	0.663684
ENSECAG00000010824	LTN1							Up704	0.663631
ENSECAG00000015791	PRIM2							Up706	0.662496
ENSECAG00000013918	ZFYVE16							Up707	0.662468
ENSECAG00000014712	DDHD1							Up708	0.661886
ENSECAG00000024185	GNG11							Up709	0.661846
ENSECAG00000021950	NPM1							Up710	0.661721
ENSECAG00000009877	HSPA4L							Up711	0.661616
ENSECAG00000017076	NSRP1							Up712	0.661435
ENSECAG00000005377	PREPL							Up714	0.66086
ENSECAG00000019340	ZNF800							Up715	0.660293
ENSECAG00000027692	COX2							Up716	0.660201
ENSECAG00000002113	BARD1							Up717	0.660082
ENSECAG00000006243	RANBP6							Up718	0.659932
ENSECAG00000020917	STAG2							Up719	0.659907
ENSECAG00000018322	EXO1							Up720	0.659594
ENSECAG00000008574	RNF138							Up721	0.659301
ENSECAG00000019522	ENSECAG00000019522							Up722	0.658949
ENSECAG00000020774	WEE1							Up723	0.658521
ENSECAG00000020044	IMPAD1							Up724	0.65851
ENSECAG00000008412	ICE2							Up725	0.658203
ENSECAG00000019898	TTC21B							Up726	0.658141
ENSECAG00000019399	TGS1							Up727	0.658129
ENSECAG00000014866	IRAK4							Up728	0.658051
ENSECAG00000014262	RNASEH2B							Up729	0.657926
ENSECAG00000017806	GGPS1							Up730	0.657738
ENSECAG00000012988	UCHL5							Up731	0.656946
ENSECAG00000005171	ENSECAG00000005171							Up732	0.656825
ENSECAG00000010468	CHCHD7							Up733	0.656635
ENSECAG00000012479	LARP4							Up734	0.656518
ENSECAG00000001911	TAX1BP1							Up735	0.656394
ENSECAG00000013304	RFWD2							Up736	0.65623
ENSECAG00000016930	ZNF713							Up737	0.656069
ENSECAG00000004675	PLEKHF2							Up738	0.656047
ENSECAG00000007540	CYP7B1							Up739	0.655746
ENSECAG00000018247	ZNF404							Up741	0.655043
ENSECAG00000005156	ZFP2							Up742	0.654878
ENSECAG00000022659	CCL26							Up743	0.654738
ENSECAG00000013200	PEX3							Up745	0.653527
ENSECAG00000009573	RPS3A							Up746	0.653215

APPENDIX 63: List of DEGs IN U/63-infected samples – Part 18

		U/63-WT		U/63-E186K		U/63-219		U/63-E186K-219	
Gene ID	Gene name	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC
ENSECAG00000020668	ENSECAG00000020668							Up747	0.652882
ENSECAG00000018957	CEBPZ							Up748	0.652833
ENSECAG00000014401	ING5							Up749	0.652805
ENSECAG00000009478	RHOT1							Up750	0.652313
ENSECAG00000017599	ZNF567							Up751	0.652185
ENSECAG00000018921	MTMR2							Up753	0.652013
ENSECAG00000016110	EXTL2							Up754	0.65193
ENSECAG00000010517	PRRG1							Up755	0.651351
ENSECAG00000013971	ATG5							Up756	0.65129
ENSECAG00000000568	TMEM68							Up757	0.650895
ENSECAG00000005128	ZNF30							Up758	0.650494
ENSECAG00000006925	WDR53							Up760	0.650061
ENSECAG00000016657	ZNF211							Up761	0.650016
ENSECAG00000017444	TAB3							Up762	0.649727
ENSECAG00000006127	SBNO1							Up763	0.649371
ENSECAG00000003569	ST6GAL1							Up764	0.649183
ENSECAG00000006287	PARP11							Up765	0.648618
ENSECAG00000022028	C9orf72							Up766	0.648482
ENSECAG00000001445	CHD7							Up767	0.648407
ENSECAG00000027681	ND2							Up768	0.6484
ENSECAG00000011405	BLM							Up771	0.648162
ENSECAG00000025074	C9orf64							Up772	0.647274
ENSECAG00000015542	OSBPL8							Up773	0.647227
ENSECAG00000011182	PUS7L							Up774	0.646034
ENSECAG00000010941	CCDC150							Up775	0.645983
ENSECAG00000009970	TIAL1							Up776	0.645357
ENSECAG00000010199	EMC2							Up777	0.645357
ENSECAG00000006803	SYTL2							Up778	0.645301
ENSECAG00000024511	MRPL1							Up779	0.645211
ENSECAG00000004770	FAM217B							Up780	0.64501
ENSECAG00000015380	EIF2B3							Up781	0.644955
ENSECAG00000004461	TIGD7							Up782	0.644872
ENSECAG00000020867	ACTR3							Up783	0.644559
ENSECAG00000014213	SRPK1							Up784	0.644481
ENSECAG00000005248	ZNF146							Up785	0.644297
ENSECAG00000014236	ENSECAG00000014236							Up786	0.644281
ENSECAG00000023777	WDR43							Up787	0.644272
ENSECAG00000011902	RAB8B							Up788	0.644066
ENSECAG00000016552	PAXBP1							Up789	0.643204
ENSECAG00000023235	KPNA4							Up790	0.642234
ENSECAG00000019161	DIAPH3							Up791	0.641917
ENSECAG00000016031	AGGF1							Up792	0.641314
ENSECAG00000000329	C6orf62							Up793	0.641242
ENSECAG00000008769	VAPA							Up794	0.641054
ENSECAG00000023450	KATNBL1							Up795	0.641049
ENSECAG00000009131	OMA1							Up796	0.640742
ENSECAG00000020320	CD302							Up797	0.64046
ENSECAG00000015379	FGD4							Up798	0.640379
ENSECAG00000000276	DNAL1							Up799	0.639485
ENSECAG00000005949	MSH6							Up800	0.638933
ENSECAG00000024208	GRSF1							Up801	0.638033
ENSECAG00000023010	USP45							Up802	0.637898
ENSECAG00000020263	ACOT13							Up803	0.636966
ENSECAG00000024919	CDKAL1							Up804	0.636863
ENSECAG00000026878	MAGOHB							Up805	0.63661
ENSECAG00000004372	SLC22A3							Up806	0.63636
ENSECAG00000000595	SLC30A9							Up808	0.635839
ENSECAG00000019885	ATP1B3							Up810	0.635221
ENSECAG00000004167	ZBED8							Up811	0.635148
ENSECAG00000019406	PDCD5							Up812	0.635069
ENSECAG00000010243	GALNT7							Up813	0.634973
ENSECAG00000002187	C1D							Up814	0.634934
ENSECAG00000010431	THOC2							Up815	0.634399
ENSECAG00000000004	RNF170							Up816	0.633852
ENSECAG00000020692	C2CD5							Up817	0.633805
ENSECAG00000014948	KIAA1524							Up818	0.633697
ENSECAG00000010253	PROKR2							Up819	0.633046
ENSECAG00000016758	ENSECAG00000016758							Up820	0.6323
ENSECAG00000007826	DONSON							Up822	0.631923
ENSECAG00000000007	TMEM70							Up823	0.631686
ENSECAG00000016155	KIF18A							Up824	0.631676
ENSECAG00000005718	MAP4K5							Up825	0.631569
ENSECAG00000010450	CCDC59							Up826	0.631421

APPENDIX 64: List of DEGs IN U/63-infected samples – Part 19

		U/63-WT		U/63-E186K		U/63-219		U/63-E186K-219	
Gene ID	Gene name	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC
ENSECAG00000018424	EIF2S1							Up827	0.631056
ENSECAG00000013355	APOPT1							Up828	0.630967
ENSECAG00000017176	EYA3							Up829	0.630855
ENSECAG00000020050	RPAP3							Up830	0.630823
ENSECAG00000016099	BMP2K							Up831	0.630087
ENSECAG00000007375	ENSECAG00000007375							Up832	0.629973
ENSECAG00000021072	KNTC1							Up833	0.629324
ENSECAG00000003237	GCC2							Up834	0.62894
ENSECAG00000014059	ECT2							Up835	0.628538
ENSECAG00000015182	OSBPL1A							Up836	0.628381
ENSECAG00000021929	DNM1L							Up837	0.627679
ENSECAG00000004593	HAUS3							Up838	0.627444
ENSECAG00000009229	APOOL							Up839	0.627375
ENSECAG00000012941	JMJD1C							Up840	0.627194
ENSECAG00000014803	GTF3A							Up841	0.626432
ENSECAG00000009060	C3orf58							Up842	0.626395
ENSECAG00000020058	PNPT1							Up843	0.626058
ENSECAG00000013328	TMOD2							Up844	0.625688
ENSECAG00000003367	RAD51C							Up846	0.624968
ENSECAG00000013284	ENSECAG00000013284							Up847	0.624319
ENSECAG00000016722	CENPU							Up848	0.624072
ENSECAG00000014597	MCTS1							Up849	0.624
ENSECAG00000014963	PIK3C2A							Up850	0.623946
ENSECAG00000019722	AKT3							Up851	0.623908
ENSECAG00000025094	TMEM161B							Up852	0.623654
ENSECAG00000019778	CFAP97							Up853	0.623572
ENSECAG00000023836	HCFC2							Up854	0.623472
ENSECAG00000021610	GPBP1							Up855	0.623467
ENSECAG00000026949	MTHFD2L							Up856	0.62339
ENSECAG00000013560	OSGIN2							Up857	0.623357
ENSECAG00000012944	UBXN8							Up858	0.622873
ENSECAG00000020143	ENSECAG00000020143							Up860	0.622502
ENSECAG00000008626	KLHL28							Up861	0.622129
ENSECAG00000011172	PPP4R3A							Up862	0.621925
ENSECAG00000000010	LRRC58							Up863	0.621897
ENSECAG00000011242	DRAM2							Up864	0.621656
ENSECAG00000021923	GTF2F2							Up865	0.621473
ENSECAG00000009675	N4BP2							Up866	0.621405
ENSECAG00000005393	ATP11C							Up867	0.621373
ENSECAG00000006861	R3HCC1L							Up868	0.620645
ENSECAG00000021921	SCAMP1							Up869	0.620496
ENSECAG00000019938	GFPT1							Up870	0.620092
ENSECAG00000006588	CHUK							Up871	0.619812
ENSECAG00000010914	CHSY1							Up872	0.619619
ENSECAG00000010426	RECK							Up873	0.619571
ENSECAG00000010332	RRP15							Up874	0.61948
ENSECAG00000018037	CD2AP							Up875	0.619333
ENSECAG00000013688	ENSECAG00000013688							Up876	0.619272
ENSECAG00000018517	MCPH1							Up877	0.619116
ENSECAG00000027694	ND5							Up878	0.618683
ENSECAG00000016939	RPS24							Up879	0.618245
ENSECAG00000011491	FAM91A1							Up880	0.618218
ENSECAG00000009435	FIGNL1							Up881	0.617947
ENSECAG00000010684	PHIP							Up882	0.617723
ENSECAG00000019980	PGAP1							Up883	0.616933
ENSECAG00000014575	RB1							Up884	0.616896
ENSECAG00000006239	NUP155							Up885	0.616852
ENSECAG00000024193	ITGB1BP1							Up886	0.616735
ENSECAG00000011526	SLC35B3							Up887	0.616675
ENSECAG00000013521	MAD2L1							Up888	0.616634
ENSECAG00000019646	AHI1							Up889	0.616496
ENSECAG00000020833	ZNF234							Up891	0.61598
ENSECAG00000014265	ARMC8							Up892	0.615515
ENSECAG00000008550	WBP11							Up893	0.615118
ENSECAG00000023517	FAM49B							Up894	0.614997
ENSECAG00000010069	XRCC2							Up895	0.614995
ENSECAG00000014503	MMADHC							Up896	0.614628
ENSECAG00000020309	RABGGTB							Up897	0.614615
ENSECAG00000021626	DBT							Up898	0.614552
ENSECAG00000017202	SNX5							Up899	0.614475
ENSECAG00000025004	NUP54							Up900	0.614233
ENSECAG00000023828	GPAT3							Up901	0.614101
ENSECAG00000023301	GPX8							Up902	0.613921

APPENDIX 65: List of DEGs IN U/63-infected samples – Part 20

		U/63-WT		U/63-E186K		U/63-219		U/63-E186K-219	
Gene ID	Gene name	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC
ENSECAG00000021005	ENSECAG00000021005							Up903	0.613895
ENSECAG00000020987	LCA5							Up904	0.613778
ENSECAG00000023731	ABCD3							Up905	0.613573
ENSECAG00000000497	THOC7							Up906	0.613469
ENSECAG00000022483	PTP4A1							Up907	0.613004
ENSECAG00000014982	CCDC112							Up908	0.61285
ENSECAG00000020406	PPP1R12A							Up909	0.612613
ENSECAG00000021849	TFPI							Up910	0.612557
ENSECAG00000018257	ENSECAG00000018257							Up911	0.612482
ENSECAG00000007715	ZNF638							Up912	0.612245
ENSECAG00000007382	FAM92A1							Up913	0.612207
ENSECAG00000019687	ENSECAG00000019687							Up914	0.611972
ENSECAG00000007737	PLEKHA8							Up915	0.611969
ENSECAG00000011795	OSGEPL1							Up916	0.611905
ENSECAG00000014704	LRRCC1							Up917	0.611726
ENSECAG00000020597	ENSECAG00000020597							Up918	0.61133
ENSECAG00000003621	PTRH2							Up919	0.611244
ENSECAG00000015601	XPR1							Up920	0.610509
ENSECAG00000019812	AGTPBP1							Up921	0.610393
ENSECAG00000024333	TAF2							Up922	0.610148
ENSECAG00000016890	FGFR1OP2							Up923	0.609807
ENSECAG00000006857	DCAF10							Up924	0.609517
ENSECAG00000015865	GLUL							Up925	0.609329
ENSECAG00000006476	LRRC8D							Up926	0.60914
ENSECAG00000013814	HECTD2							Up927	0.609073
ENSECAG00000016487	ATP6V1E1							Up928	0.608788
ENSECAG00000006869	MAP9							Up929	0.608491
ENSECAG00000012526	DNAJB9							Up930	0.608325
ENSECAG00000015965	IPO7							Up931	0.608282
ENSECAG00000000423	USP16							Up932	0.608197
ENSECAG00000016394	IL1R1							Up933	0.607954
ENSECAG00000011689	SLBP							Up934	0.607928
ENSECAG00000017780	TRPM7							Up935	0.607613
ENSECAG00000023470	ATL2							Up936	0.607602
ENSECAG00000012679	RIF1							Up937	0.607193
ENSECAG00000014912	CCDC91							Up938	0.607192
ENSECAG00000018695	ENSECAG00000018695							Up939	0.607027
ENSECAG00000021751	ZFP14							Up940	0.607022
ENSECAG00000026902	MAP4K3							Up941	0.606687
ENSECAG00000012298	ENSECAG00000012298							Up942	0.606104
ENSECAG00000021067	ENSECAG00000021067							Up943	0.606021
ENSECAG00000018145	NT5DC3							Up944	0.605246
ENSECAG00000013853	CCDC90B							Up945	0.605135
ENSECAG00000017131	ENSECAG00000017131							Up946	0.604861
ENSECAG00000010204	ERI2							Up947	0.604669
ENSECAG00000024473	MCM10							Up948	0.604605
ENSECAG00000021804	PELI1							Up949	0.604562
ENSECAG00000022755	EIF5A2							Up950	0.603978
ENSECAG00000010712	AP3S1							Up951	0.603814
ENSECAG00000010545	TTC39C							Up952	0.603787
ENSECAG00000002701	UHRF1BP1L							Up953	0.603673
ENSECAG00000022985	USP14							Up955	0.603344
ENSECAG00000012286	DDX52							Up956	0.603277
ENSECAG00000004792	UBE3A							Up957	0.603149
ENSECAG00000000259	RTKN2							Up958	0.602764
ENSECAG00000012255	MTFMT							Up961	0.602221
ENSECAG00000014564	UFM1							Up960	0.602221
ENSECAG00000010212	SLK							Up962	0.60161
ENSECAG00000011393	GKAP1							Up963	0.601583
ENSECAG00000015913	ENSECAG00000015913							Up964	0.601426
ENSECAG00000014922	EPS8							Up966	0.601129
ENSECAG00000005271	FASTKD5							Up968	0.600919
ENSECAG00000012742	CXCL6							Up969	0.600719
ENSECAG00000021752	FBXO33							Up970	0.600492
ENSECAG00000024758	TMX3							Up971	0.600366
ENSECAG00000007182	ORC6							Up972	0.600128
ENSECAG00000020698	MELK							Up973	0.600067
ENSECAG00000010710	WDR75							Up974	0.600004
ENSECAG00000010756	PPME1							Up975	0.599442
ENSECAG00000024870	CWH43							Up976	0.598294
ENSECAG00000011778	C11orf58							Up977	0.598282
ENSECAG00000024188	ANKRD49							Up978	0.598259
ENSECAG00000021389	NCKAP1							Up979	0.598013

APPENDIX 66: List of DEGs IN U/63-infected samples – Part 21

		U/63-WT		U/63-E186K		U/63-219		U/63-E186K-219	
Gene ID	Gene name	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC	Rank	Log2FC
ENSECAG00000000513	PRELID3B							Up980	0.597944
ENSECAG00000014838	TSC22D1							Up981	0.597818
ENSECAG00000023656	DCAF17							Up983	0.597369
ENSECAG00000019466	ARIH1							Up984	0.597355
ENSECAG00000020790	ABHD17B							Up985	0.597056
ENSECAG00000010493	ENTPD4							Up986	0.59561
ENSECAG00000020166	YEATS4							Up987	0.595543
ENSECAG00000025100	DHX36							Up988	0.59525
ENSECAG00000019292	TCAIM							Up990	0.594953
ENSECAG00000011846	RTN4							Up991	0.594947
ENSECAG00000010089	DUSP6							Up992	0.594925
ENSECAG00000011292	MSR1							Up993	0.594791
ENSECAG00000008462	RYK							Up994	0.594177
ENSECAG00000008481	ENSECAG00000008481							Up995	0.593958
ENSECAG00000018362	ENSECAG00000018362							Up996	0.593942
ENSECAG00000002193	ENSECAG00000002193							Up997	0.593905
ENSECAG00000022063	ABHD5							Up998	0.593789
ENSECAG00000017558	MLKL							Up999	0.593332
ENSECAG00000014592	WDR45B							Up1000	0.593062
ENSECAG00000009788	TBC1D31							Up1001	0.59305
ENSECAG00000021819	PPIL4							Up1002	0.592949
ENSECAG00000021498	SNX13							Up1003	0.592929
ENSECAG00000019035	LIPT1							Up1004	0.592871
ENSECAG00000013610	POLR2B							Up1005	0.592659
ENSECAG00000020223	HACD1							Up1006	0.592448
ENSECAG00000008861	RRN3							Up1007	0.592393
ENSECAG00000008450	ATP2B1							Up1008	0.591835
ENSECAG00000018048	TRPC1							Up1009	0.591816
ENSECAG00000016522	SCG3							Up1010	0.591498
ENSECAG00000009904	NDUFV2							Up1011	0.591493
ENSECAG00000007655	QSER1							Up1012	0.59143
ENSECAG00000005083	IL1RAP							Up1013	0.591414
ENSECAG00000000048	UBA2							Up1014	0.591173
ENSECAG00000012273	UBA6							Up1015	0.591047
ENSECAG00000017880	SEH1L							Up1016	0.590664
ENSECAG00000009920	POLR3F							Up1017	0.590595
ENSECAG00000008506	HMMR							Up1018	0.590134
ENSECAG00000010292	NKIRAS1							Up1019	0.589681
ENSECAG00000009947	TTC33							Up1021	0.589129
ENSECAG00000014071	SLC30A7							Up1022	0.589024
ENSECAG00000014116	SLC38A9							Up1023	0.588865
ENSECAG00000000689	PPP4R2							Up1024	0.588642
ENSECAG00000026842	ACER3							Up1025	0.588102
ENSECAG00000000236	EIF3E							Up1026	0.588082
ENSECAG00000009913	CDC6							Up1027	0.587068
ENSECAG00000013019	CCDC34							Up1028	0.586869
ENSECAG00000008167	SPPL2A							Up1029	0.586726
ENSECAG00000018948	HSP90AA1							Up1030	0.586602
ENSECAG00000012605	RNGTT							Up1032	0.586522
ENSECAG00000019312	PDS5B							Up1033	0.586472
ENSECAG00000008062	DDHD2							Up1034	0.586289
ENSECAG00000009297	PLOD2							Up1035	0.585998
ENSECAG00000025163	MAP2K4							Up1036	0.585967
ENSECAG00000013510	HSPA8							Up1037	0.584972
ENSECAG00000018653	TROVE2							Up1038	0.584686
ENSECAG00000005033	WDR5B							Up1039	0.58465
ENSECAG00000009463	CCDC15							Up1040	0.584646
ENSECAG00000020103	DSTN							Up1041	0.584595
ENSECAG00000024547	ASNSD1							Up1042	0.584521
ENSECAG00000011505	CHD9							Up1043	0.584445
ENSECAG00000021993	GCLM							Up1044	0.584316
ENSECAG00000019120	ENSECAG00000019120							Up1045	0.584161
ENSECAG00000016982	CCT4							Up1046	0.583786
ENSECAG00000020616	TMTC3							Up1047	0.583738
ENSECAG00000003331	PLEKHH2							Up1048	0.583324
ENSECAG00000010790	COG6							Up1049	0.583296
ENSECAG00000005374	WAC							Up1050	0.58319
ENSECAG00000000511	CFL2							Up1051	0.583148
ENSECAG00000020362	MND1							Up1052	0.583124
ENSECAG00000007763	ALMS1							Up1053	0.582852
ENSECAG00000020945	MNS1							Up1054	0.582136
ENSECAG00000000018	KANSL1L							Up1055	0.581878
ENSECAG00000020046	PSMD14							Up1056	0.581702

APPENDIX 67: List of DEGs IN U/63-infected samples – Part 22

		U/63-WT		U/63-E186K		U/63-219		U/63-E186K-219	
Gene ID	**Gene name**	**Rank**	**Log2FC**	**Rank**	**Log2FC**	**Rank**	**Log2FC**	**Rank**	**Log2FC**
ENSECAG00000021982	PRKRIR							Up1057	0.581416
ENSECAG00000024474	SLC30A5							Up1058	0.581388
ENSECAG00000013713	FAM45A							Up1059	0.581057
ENSECAG00000004843	MDM4							Up1060	0.580478
ENSECAG00000020786	C1QTNF2							Down66	-0.582546
ENSECAG00000018835	FAM46B							Down65	-0.582742
ENSECAG00000021399	FAM98C							Down62	-0.588967
ENSECAG00000024890	CCDC120							Down55	-0.613613
ENSECAG00000011219	CDKN2D							Down50	-0.637784
ENSECAG00000018161	C1QTNF5							Down48	-0.647381
ENSECAG00000024202	XKR8							Down44	-0.666767
ENSECAG00000012364	MYEOV2							Down42	-0.668985
ENSECAG00000013721	SH2D5							Down40	-0.681721
ENSECAG00000006208	MBLAC1							Down36	-0.695815
ENSECAG00000014275	DBP							Down26	-0.733667
ENSECAG00000004743	PPP1R3D							Down17	-0.882141
ENSECAG00000017304	MAPK8IP2							Down16	-0.889524
ENSECAG00000006591	ATG9B							Down14	-0.895217
ENSECAG00000018855	EP400NL							Down12	-0.947197
ENSECAG00000007520	BBS5							Down8	-1.05334
ENSECAG00000002197	ENSECAG00000002197							Down6	-1.09455
ENSECAG00000013952	CADM4							Down5	-1.14208
ENSECAG00000016293	NANOS3							Down2	-1.51865

	U/63-WT	U/63-E186K	U/63-219	U/63-E186K-219
Up-regulated genes	140	146	117	1060
Down-regulated genes	258	59	99	66
Total DEGs	398	205	216	1126

REFERENCES

References

ABLASSER, A., SCHMID-BURGK, J. L., HEMMERLING, I., HORVATH, G. L., SCHMIDT, T., LATZ, E. & HORNUNG, V. 2013. Cell intrinsic immunity spreads to bystander cells via the intercellular transfer of cGAMP. Nature, 503, 530-4.

AL-GARAWI, A., HUSAIN, M., ILIEVA, D., HUMBLES, A. A., KOLBECK, R., STAMPFLI, M. R., O'BYRNE, P. M., COYLE, A. J. & JORDANA, M. 2012. Shifting of immune responsiveness to house dust mite by influenza A infection: genomic insights. J Immunol, 188, 832-43.

ALLEN, I. C., SCULL, M. A., MOORE, C. B., HOLL, E. K., MCELVANIA-TEKIPPE, E., TAXMAN, D. J., GUTHRIE, E. H., PICKLES, R. J. & TING, J. P. 2009. The NLRP3 inflammasome mediates in vivo innate immunity to influenza A virus through recognition of viral RNA. *Immunity,* 30, 556-65.

ALONSO-CAPLEN, F. V. & KRUG, R. M. 1991. Regulation of the extent of splicing of influenza virus NS1 mRNA: role of the rates of splicing and of the nucleocytoplasmic transport of NS1 mRNA. *Mol Cell Biol,* 11, 1092-8.

ALONSO-CAPLEN, F. V., NEMEROFF, M. E., QIU, Y. & KRUG, R. M. 1992. Nucleocytoplasmic transport: the influenza virus NS1 protein regulates the transport of spliced NS2 mRNA and its precursor NS1 mRNA. *Genes Dev,* 6, 255-67.

ALVES BEUTTEMMULLER, E., WOODWARD, A., RASH, A., DOS SANTOS FERRAZ, L. E., FERNANDES ALFIERI, A., ALFIERI, A. A. & ELTON, D. 2016. Characterisation of the epidemic strain of H3N8 equine influenza virus responsible for outbreaks in South America in 2012. *Virol J,* 13, 45.

ANASTASINA, M., LE MAY, N., BUGAI, A., FU, Y., SODERHOLM, S., GAELINGS, L., OHMAN, T., TYNELL, J., KYTTANEN, S., BARBORIC, M., NYMAN, T. A., MATIKAINEN, S., JULKUNEN, I., BUTCHER, S. J., EGLY, J. M. & KAINOV, D. E. 2016. Influenza virus NS1 protein binds cellular DNA to block transcription of antiviral genes. Biochim Biophys Acta, 1859, 1440-1448.

AOKI, M., AOKI, H., RAMANATHAN, R., HAIT, N. C. & TAKABE, K. 2016. Sphingosine-1-Phosphate Signaling in Immune Cells and Inflammation: Roles and Therapeutic Potential. Mediators Inflamm, 2016, 8606878.

ARAGON, T., DE LA LUNA, S., NOVOA, I., CARRASCO, L., ORTIN, J. & NIETO, A. 2000. Eukaryotic translation initiation factor 4GI is a cellular target for NS1 protein, a translational activator of influenza virus. *Mol Cell Biol,* 20, 6259-68.

ARAMINI, J. M., MA, L. C., ZHOU, L., SCHAUDER, C. M., HAMILTON, K., AMER, B. R., MACK, T. R., LEE, H. W., CICCOSANTI, C. T., ZHAO, L., XIAO, R., KRUG, R. M. & MONTELIONE, G. T. 2011. Dimer interface of the effector domain of non-structural protein 1 from influenza A virus: an interface with multiple functions. J Biol Chem, 286, 26050-60.

ARTHUR, R. J. & SUANN, C. J. 2011. Biosecurity and vaccination strategies to minimise the effect of an equine influenza outbreak on racing and breeding. *Aust Vet J,* 89 Suppl 1, 109-13.

ARUGA, J., YOKOTA, N. & MIKOSHIBA, K. 2003. Human SLITRK family genes: genomic organization and expression profiling in normal brain and brain tumor tissue. Gene, 315, 87-94.

AYLLON, J. & GARCIA-SASTRE, A. 2015. The NS1 protein: a multitasking virulence factor. Curr Top Microbiol Immunol, 386, 73-107.

AYLLON, J., RUSSELL, R. J., GARCIA-SASTRE, A. & HALE, B. G. 2012. Contribution of NS1 effector domain dimerization to influenza A virus replication and virulence. J Virol, 86, 13095-8.

BALASURIYA, U. B., ZHANG, J., GO, Y. Y. & MACLACHLAN, N. J. 2014. Experiences with infectious cDNA clones of equine arteritis virus: lessons learned and insights gained. Virology, 462-463, 388-403.

BANNINGER, G. & REICH, N. C. 2004. STAT2 nuclear trafficking. *J Biol Chem,* 279**,** 39199-206.

BAO, G., XU, L., XU, X., ZHAI, L., DUAN, C., XU, D., SONG, J., LIU, Z., TAO, R., CUI, Z. & YANG, H. 2016. SGTB Promotes the Caspase-Dependent Apoptosis in Chondrocytes of Osteoarthritis. Inflammation, 39, 601-10.

BARABINO, S. M., HUBNER, W., JENNY, A., MINVIELLE-SEBASTIA, L. & KELLER, W. 1997. The 30-kD subunit of mammalian cleavage and polyadenylation specificity factor and its yeast homolog are RNA-binding zinc finger proteins. *Genes Dev,* 11**,** 1703-16.

BARBA, M. & DALY, J. M. 2016. The Influenza NS1 Protein: What Do We Know in Equine Influenza Virus Pathogenesis? Pathogens, 5(3). Pii: E57.

BASLER, C. F., REID, A. H., DYBING, J. K., JANCZEWSKI, T. A., FANNING, T. G., ZHENG, H., SALVATORE, M., PERDUE, M. L., SWAYNE, D. E., GARCIA-SASTRE, A., PALESE, P. & TAUBENBERGER, J. K. 2001. Sequence of the 1918 pandemic influenza virus nonstructural gene (NS) segment and characterization of recombinant viruses bearing the 1918 NS genes. *Proc Natl Acad Sci U S A,* 98**,** 2746-51.

BAVAGNOLI, L., CUCUZZA, S., CAMPANINI, G., ROVIDA, F., PAOLUCCI, S., BALDANTI, F. & MAGA, G. 2015. The novel influenza A virus protein PA-X and its naturally deleted variant show different enzymatic properties in comparison to the viral endonuclease PA. Nucleic Acids Res, 43, 9405-17.
BERGMANN, M., GARCIA-SASTRE, A., CARNERO, E., PEHAMBERGER, H., WOLFF, K., PALESE, P. & MUSTER, T. 2000. Influenza virus NS1 protein counteracts PKR-mediated inhibition of replication. J Virol, 74, 6203-6.

BERKOWITZ, B., HUANG, D. B., CHEN-PARK, F. E., SIGLER, P. B. & GHOSH, G. 2002. The x-ray crystal structure of the NF-kappa B p50.p65 heterodimer bound to the interferon beta -kappa B site. *J Biol Chem,* 277**,** 24694-700.

BERRIOS, P. 2005. [Equine influenza in Chile (1963-1992): a possible human case]. *Rev Chilena Infectol,* 22**,** 47-50.

BODIN, L., VERSTUYFT, C., TREGOUET, D. A., ROBERT, A., DUBERT, L., FUNCK-BRENTANO, C., JAILLON, P., BEAUNE, P., LAURENT-PUIG, P., BECQUEMONT, L. & LORIOT, M. A. 2005. Cytochrome P450 2C9 (CYP2C9) and vitamin K epoxide reductase (VKORC1) genotypes as determinants of acenocoumarol sensitivity. Blood, 106, 135-40.

BOGUNOVIC, D., BYUN, M., DURFEE, L. A., ABHYANKAR, A., SANAL, O., MANSOURI, D., SALEM, S., RADOVANOVIC, I., GRANT, A. V., ADIMI, P., MANSOURI, N., OKADA, S., BRYANT, V. L., KONG, X. F., KREINS, A., VELEZ, M. M., BOISSON, B., KHALILZADEH, S., OZCELIK, U., DARAZAM, I. A., SCHOGGINS, J. W., RICE, C. M., AL-MUHSEN, S., BEHR, M., VOGT, G., PUEL, A., BUSTAMANTE, J., GROS, P., HUIBREGTSE, J. M., ABEL, L., BOISSON-DUPUIS, S. & CASANOVA, J. L. 2012. Mycobacterial disease and impaired IFN-gamma immunity in humans with inherited ISG15 deficiency. *Science,* 337**,** 1684-8.

BORNHOLDT, Z. A. & PRASAD, B. V. 2006. X-ray structure of influenza virus NS1 effector domain. *Nat Struct Mol Biol,* 13**,** 559-60.

BRACIALE, T. J., SUN, J. & KIM, T. S. 2012. Regulating the adaptive immune response to respiratory virus infection. *Nat Rev Immunol,* 12**,** 295-305.

BRAZIL, D. P., YANG, Z. Z. & HEMMINGS, B. A. 2004. Advances in protein kinase B signalling: AKTion on multiple fronts. Trends Biochem Sci, 29, 233-42.

BRIDEAU-ANDERSEN, A. D., HUANG, X., SUN, S. C., CHEN, T. T., STARK, D., SAS, I. J., ZADIK, L., DAWES, G. N., GUPTILL, D. R., MCCORD, R., GOVINDARAJAN, S., ROY, A., YANG, S., GAO, J., CHEN, Y. H., SKARTVED, N. J., PEDERSEN, A. K., LIN, D., LOCHER, C. P., REBBAPRAGADA, I., JENSEN, A. D., BASS, S. H., NISSEN, T. L., VISWANATHAN, S., FOSTER, G. R., SYMONS, J. A. & PATTEN, P. A.

2007. Directed evolution of gene-shuffled IFN-alpha molecules with activity profiles tailored for treatment of chronic viral diseases. *Proc Natl Acad Sci U S A,* 104**,** 8269-74.
BROWN, E. G., LIU, H., KIT, L. C., BAIRD, S. & NESRALLAH, M. 2001. Pattern of mutation in the genome of influenza A virus on adaptation to increased virulence in the mouse lung: identification of functional themes. Proc Natl Acad Sci U S A, 98, 6883-8.

BRYANT, N. A., PAILLOT, R., RASH, A. S., MEDCALF, E., MONTESSO, F., ROSS, J., WATSON, J., JEGGO, M., LEWIS, N. S., NEWTON, J. R. & ELTON, D. M. 2010. Comparison of two modern vaccines and previous influenza infection against challenge with an equine influenza virus from the Australian 2007 outbreak. *Vet Res,* 41**,** 19.

BRYANT, N. A., RASH, A. S., RUSSELL, C. A., ROSS, J., COOKE, A., BOWMAN, S., MACRAE, S., LEWIS, N. S., PAILLOT, R., ZANONI, R., MEIER, H., GRIFFITHS, L. A., DALY, J. M., TIWARI, A., CHAMBERS, T. M., NEWTON, J. R. & ELTON, D. M. 2009. Antigenic and genetic variations in European and North American equine influenza virus strains (H3N8) isolated from 2006 to 2007. *Vet Microbiol,* 138**,** 41-52.

BRYANT, N. A., RASH, A. S., WOODWARD, A. L., MEDCALF, E., HELWEGEN, M., WOHLFENDER, F., CRUZ, F., HERRMANN, C., BORCHERS, K., TIWARI, A., CHAMBERS, T. M., NEWTON, J. R., MUMFORD, J. A. & ELTON, D. M. 2011. Isolation and characterisation of equine influenza viruses (H3N8) from Europe and North America from 2008 to 2009. *Vet Microbiol,* 147**,** 19-27.

BUDD, M. E. & CAMPBELL, J. L. 2013. Dna2 is involved in CA strand resection and nascent lagging strand completion at native yeast telomeres. J Biol Chem, 288, 29414-29.

BURGUI, I., ARAGON, T., ORTIN, J. & NIETO, A. 2003. PABP1 and eIF4GI associate with influenza virus NS1 protein in viral mRNA translation initiation complexes. *J Gen Virol,* 84**,** 3263-74.

BURNELL, F. J., HOLMES, M. A., ROIKO, A. H., LOWE, J. B., HEIL, G. L., WHITE, S. K. & GRAY, G. C. 2014. Little evidence of human infection with equine influenza during the 2007 epizootic, Queensland, Australia. *J Clin Virol,* 59**,** 100-3.

BURROWS, R., DENYER, M., GOODRIDGE, D. & HAMILTON, F. 1981. Field and laboratory studies of equine influenza viruses isolated in 1979. *Vet Rec,* 109**,** 353-6.

CAL, S., OBAYA, A. J., LLAMAZARES, M., GARABAYA, C., QUESADA, V. & LOPEZ-OTIN, C. 2002. Cloning, expression analysis, and structural characterization of seven novel human ADAMTSs, a family of metalloproteinases with disintegrin and thrombospondin-1 domains. Gene, 283, 49-62.
CANTLEY, L. C. 2002. The phosphoinositide 3-kinase pathway. Science, 296, 1655-7.

CARRILLO, B., CHOI, J. M., BORNHOLDT, Z. A., SANKARAN, B., RICE, A. P. & PRASAD, B. V. 2014. The influenza A virus protein NS1 displays structural polymorphism. *J Virol,* 88**,** 4113-22.

CATALDO, D. D., GUEDERS, M. M., ROCKS, N., SOUNNI, N. E., EVRARD, B., BARTSCH, P., LOUIS, R., NOEL, A. & FOIDART, J. M. 2003. Pathogenic role of matrix metalloproteases and their inhibitors in asthma and chronic obstructive pulmonary disease and therapeutic relevance of matrix metalloproteases inhibitors. Cell Mol Biol (Noisy-le-grand), 49, 875-84.

CAUTHEN, A. N., SWAYNE, D. E., SEKELLICK, M. J., MARCUS, P. I. & SUAREZ, D. L. 2007. Amelioration of influenza virus pathogenesis in chickens attributed to the enhanced interferon-inducing capacity of a virus with a truncated NS1 gene. J Virol, 81, 1838-47.

CHAN, Y. W. & WEST, S. 2015. GEN1 promotes Holliday junction resolution by a coordinated nick and counter-nick mechanism. Nucleic Acids Res, 43, 10882-92.

CHAIPAN, C., KOBASA, D., BERTRAM, S., GLOWACKA, I., STEFFEN, I., TSEGAYE, T.S., TAKEDA, M., BUGGE, T.H., KIM, S., PARK, Y., MARZI, A., POHLMANN, S. 2009. Proteolytic Activation of the 1918 Influenza Virus Hemagglutinin. J Virol, 83, 3200-3211.

CHEN, G., LIU, C. H., ZHOU, L. & KRUG, R. M. 2014. Cellular DDX21 RNA helicase inhibits influenza A virus replication but is counteracted by the viral NS1 protein. *Cell Host Microbe,* 15**,** 484-93.

CHEN, L. & FLIES, D. B. 2013. Molecular mechanisms of T cell co-stimulation and co-inhibition. Nat Rev Immunol, 13, 227-42.

CHEN, M. H. & MALBON, C. C. 2009. G-protein-coupled receptor-associated A-kinase anchoring proteins AKAP5 and AKAP12: differential trafficking and distribution. Cell Signal, 21, 136-42.
CHEN, N., WANG, S., SMENTEK, L., HESS, B. A., JR. & WU, R. 2015. Biosynthetic Mechanism of Lanosterol: Cyclization. Angew Chem Int Ed Engl, 54, 8693-6.

CHEN, Y., WANG, S., YI, Z., TIAN, H., ALIYARI, R., LI, Y., CHEN, G., LIU, P., ZHONG, J., CHEN, X., DU, P., SU, L., QIN, F. X., DENG, H. & CHENG, G. 2014. Interferon-inducible cholesterol-25-hydroxylase inhibits hepatitis C virus replication via distinct mechanisms. Sci Rep, 4, 7242.

CHEN, Z., LI, Y. & KRUG, R. M. 1999. Influenza A virus NS1 protein targets poly(A)-binding protein II of the cellular 3'-end processing machinery. EMBO J, 18, 2273-83.

CHEUNG, C. Y., POON, L. L., LAU, A. S., LUK, W., LAU, Y. L., SHORTRIDGE, K. F., GORDON, S., GUAN, Y. & PEIRIS, J. S. 2002. Induction of proinflammatory cytokines in human macrophages by influenza A (H5N1) viruses: a mechanism for the unusual severity of human disease? *Lancet,* 360**,** 1831-7.

CHEUNG, D. H., KUNG, H. F., HUANG, J. J. & SHAW, P. C. 2012. PinX1 is involved in telomerase recruitment and regulates telomerase function by mediating its localization. FEBS Lett, 586, 3166-71.

CHI, H. 2011. Sphingosine-1-phosphate and immune regulation: trafficking and beyond. Trends Pharmacol Sci, 32, 16-24.

CHIAUR, D. S., MURTHY, S., CENCIARELLI, C., PARKS, W., LODA, M., INGHIRAMI, G., DEMETRICK, D. & PAGANO, M. 2000. Five human genes encoding F-box proteins: chromosome mapping and analysis in human tumors. Cytogenet Cell Genet, 88, 255-8.

CHIEN, C. Y., TEJERO, R., HUANG, Y., ZIMMERMAN, D. E., RIOS, C. B., KRUG, R. M. & MONTELIONE, G. T. 1997. A novel RNA-binding motif in influenza A virus non-structural protein 1. *Nat Struct Biol,* 4**,** 891-5.

CHIEN, C. Y., XU, Y., XIAO, R., ARAMINI, J. M., SAHASRABUDHE, P. V., KRUG, R. M. & MONTELIONE, G. T. 2004. Biophysical characterization of the complex between double-stranded RNA and the N-terminal domain of the NS1 protein from influenza A virus: evidence for a novel RNA-binding mode. *Biochemistry,* 43**,** 1950-62.

CHILDS, K. S., RANDALL, R. E. & GOODBOURN, S. 2013. LGP2 plays a critical role in sensitizing mda-5 to activation by double-stranded RNA. *PLoS One,* 8**,** e64202.

CHINI, C. C. & CHEN, J. 2003. Human claspin is required for replication checkpoint control. J Biol Chem, 278, 30057-62.

CLARK, A. M., NOGALES, A., MARTINEZ-SOBRIDO, L., TOPHAM, D. J. & DEDIEGO, M. L. 2017. Functional Evolution of Influenza Virus NS1 Protein in Currently Circulating Human 2009 Pandemic H1N1 Viruses. J Virol, 91(17). Pii: e00721-17.

COBURN 2017. Molecular determinants of Influenza A virus cross-species jumps. PhD Thesis.

COLGAN, D. F. & MANLEY, J. L. 1997. Mechanism and regulation of mRNA polyadenylation. *Genes Dev,* 11**,** 2755-66.

CONESA, A., MADRIGAL, P., TARAZONA, S., GOMEZ-CABRERO, D., CERVERA, A., MCPHERSON, A., SZCZESNIAK, M. W., GAFFNEY, D. J., ELO, L. L., ZHANG, X. & MORTAZAVI, A. 2016. A survey of best practices for RNA-seq data analysis. Genome Biol, 17, 13.

CONSORTIUM, G. T. 2013. The Genotype-Tissue Expression (GTEx) project. Nat Genet, 45, 580-5.

CONTI, L., SIPIONE, S., MAGRASSI, L., BONFANTI, L., RIGAMONTI, D., PETTIROSSI, V., PESCHANSKI, M., HADDAD, B., PELICCI, P., MILANESI, G., PELICCI, G. & CATTANEO, E. 2001. Shc signaling in differentiating neural progenitor cells. Nat Neurosci, 4, 579-86.

CORRIGALL, V. M., SOLAU-GERVAIS, E. & PANAYI, G. S. 2000. Lack of CD80 expression by fibroblast-like synoviocytes leading to anergy in T lymphocytes. Arthritis Rheum, 43, 1606-15.

COWLED, B., WARD, M. P., HAMILTON, S. & GARNER, G. 2009. The equine influenza epidemic in Australia: spatial and temporal descriptive analyses of a large propagating epidemic. *Prev Vet Med,* 92, 60-70.

CRAWFORD, P. C., DUBOVI, E. J., CASTLEMAN, W. L., STEPHENSON, I., GIBBS, E. P., CHEN, L., SMITH, C., HILL, R. C., FERRO, P., POMPEY, J., BRIGHT, R. A., MEDINA, M. J., JOHNSON, C. M., OLSEN, C. W., COX, N. J., KLIMOV, A. I., KATZ, J. M. & DONIS, R. O. 2005. Transmission of equine influenza virus to dogs. *Science,* 310, 482-5.

CROSS, N. A., CHANDRASEKHARAN, S., JOKONYA, N., FOWLES, A., HAMDY, F. C., BUTTLE, D. J. & EATON, C. L. 2005. The expression and regulation of ADAMTS-1, -4, -5, -9, and -15, and TIMP-3 by TGFbeta1 in prostate cells: relevance to the accumulation of versican. Prostate, 63, 269-75.

CROTTA, S., DAVIDSON, S., MAHLAKOIV, T., DESMET, C. J., BUCKWALTER, M. R., ALBERT, M. L., STAEHELI, P. & WACK, A. 2013. Type I and type III interferons drive redundant amplification loops to induce a transcriptional signature in influenza-infected airway epithelia. *PLoS Pathog,* 9, e1003773.

CRUSE, G., METCALFE, D. D. & OLIVERA, A. 2014. Functional deregulation of KIT: link to mast cell proliferative diseases and other neoplasms. Immunol Allergy Clin North Am, 34, 219-37.

DAINES, M. O., TABATA, Y., WALKER, B. A., CHEN, W., WARRIER, M. R., BASU, S. & HERSHEY, G. K. 2006. Level of expression of IL-13R alpha 2 impacts receptor distribution and IL-13 signaling. J Immunol, 176, 7495-501.

DALY, J. M., BLUNDEN, A. S., MACRAE, S., MILLER, J., BOWMAN, S. J., KOLODZIEJEK, J., NOWOTNY, N. & SMITH, K. C. 2008. Transmission of equine influenza virus to English foxhounds. *Emerg Infect Dis,* 14, 461-4.

DALY, J. M., LAI, A. C., BINNS, M. M., CHAMBERS, T. M., BARRANDEGUY, M. & MUMFORD, J. A. 1996. Antigenic and genetic evolution of equine H3N8 influenza A viruses. J Gen Virol, 77 (Pt 4), 661-71.

DALY, J. M., MACRAE, S., NEWTON, J. R., WATTRANG, E. & ELTON, D. M. 2011. Equine influenza: a review of an unpredictable virus. Vet J, 189, 7-14.

DAMIANI, A. M., SCICLUNA, M. T., CIABATTI, I., CARDETI, G., SALA, M., VULCANO, G., CORDIOLI, P., MARTELLA, V., AMADDEO, D. & AUTORINO, G. L. 2008. Genetic characterization of equine influenza viruses isolated in Italy between 1999 and 2005. *Virus Res,* 131, 100-5.

DANKAR, S. K., MIRANDA, E., FORBES, N. E., PELCHAT, M., TAVASSOLI, A., SELMAN, M., PING, J., JIA, J. & BROWN, E. G. 2013. Influenza A/Hong Kong/156/1997(H5N1) virus NS1 gene mutations F103L and M106I both increase IFN antagonism, virulence and cytoplasmic localization but differ in binding to RIG-I and CPSF30. Virol J, 10, 243.

DANKAR, S. K., WANG, S., PING, J., FORBES, N. E., KELETA, L., LI, Y. & BROWN, E. G. 2011. Influenza A virus NS1 gene mutations F103L and M106I increase replication and virulence. Virol J, 8, 13.

DAS, A., DINH, P. X., PANDA, D. & PATTNAIK, A. K. 2014. Interferon-inducible protein IFI35 negatively regulates RIG-I antiviral signaling and supports vesicular stomatitis virus replication. J Virol, 88, 3103-13.

DAS, K., MA, L. C., XIAO, R., RADVANSKY, B., ARAMINI, J., ZHAO, L., MARKLUND, J., KUO, R. L., TWU, K. Y., ARNOLD, E., KRUG, R. M. & MONTELIONE, G. T. 2008. Structural basis for suppression of a host antiviral response by influenza A virus. Proc Natl Acad Sci U S A, 105, 13093-8.

DATTA, A., GHATAK, D., DAS, S., BANERJEE, T., PAUL, A., BUTTI, R., GORAIN, M., GHUWALEWALA, S., ROYCHOWDHURY, A., ALAM, S. K., DAS, P., CHATTERJEE, R., DASGUPTA, M., PANDA, C. K., KUNDU, G. C. & ROYCHOUDHURY, S. 2017. p53 gain-of-function mutations increase Cdc7-dependent replication initiation. EMBO Rep. 18, 2030-2050.

DE LA LUNA, S., FORTES, P., BELOSO, A. & ORTIN, J. 1995. Influenza virus NS1 protein enhances the rate of translation initiation of viral mRNAs. J Virol, 69, 2427-33.

DE VRIES, W., HAASNOOT, J., FOUCHIER, R., DE HAAN, P. & BERKHOUT, B. 2009. Differential RNA silencing suppression activity of NS1 proteins from different influenza A virus strains. J Gen Virol, 90, 1916-22.

DEDIEGO, M. L., NOGALES, A., LAMBERT-EMO, K., MARTINEZ-SOBRIDO, L. & TOPHAM, D. J. 2016. NS1 Protein Mutation I64T Affects Interferon Responses and Virulence of Circulating H3N2 Human Influenza A Viruses. J Virol, 90, 9693-9711.

DETOURNAY, O., MORRISON, D.A., WAGNER, B., ZARNEGAR, B., WATTRANG, E. 2013. Genomic analysis and mRNA expression of equine type I interferon genes. J Interferon Cytokine Res, 33, 746-759.

DIEBOLD, S. S., MASSACRIER, C., AKIRA, S., PATUREL, C., MOREL, Y. & REIS E SOUSA, C. 2006. Nucleic acid agonists for Toll-like receptor 7 are defined by the presence of uridine ribonucleotides. *Eur J Immunol,* 36**,** 3256-67.

DLAMINI, Z., RUPNARAIN, C., NAICKER, S., HULL, R. & MBITA, Z. 2016. Expression analysis and association of RBBP6 with apoptosis in colon cancers. J Mol Histol, 47, 169-82.

DONATELLI, I., CASTRUCCI, M. R., DE MARCO, M. A., DELOGU, M. & WEBSTER, R. G. 2017. Human-Animal Interface: The Case for Influenza Interspecies Transmission. Adv Exp Med Biol, 972, 17-33.

DONELAN, N. R., DAUBER, B., WANG, X., BASLER, C. F., WOLFF, T. & GARCIA-SASTRE, A. 2004. The N- and C-terminal domains of the NS1 protein of influenza B virus can independently inhibit IRF-3 and beta interferon promoter activation. *J Virol,* 78**,** 11574-82.

DUFFY, S., SHACKELTON, L. A. & HOLMES, E. C. 2008. Rates of evolutionary change in viruses: patterns and determinants. Nat Rev Genet, 9, 267-76.

DUNDON, W. G. & CAPUA, I. 2009. A Closer Look at the NS1 of Influenza Virus. Viruses, 1, 1057-72.

DURFEE, L. A. & HUIBREGTSE, J. M. 2010. Identification and Validation of ISG15 Target Proteins. *Subcell Biochem,* 54**,** 228-37.

EGOROV, A., BRANDT, S., SEREINIG, S., ROMANOVA, J., FERKO, B., KATINGER, D., GRASSAUER, A., ALEXANDROVA, G., KATINGER, H. & MUSTER, T. 1998. Transfectant influenza A viruses with long deletions in the NS1 protein grow efficiently in Vero cells. J Virol, 72, 6437-41.

EHRHARDT, C. & LUDWIG, S. 2009. A new player in a deadly game: influenza viruses and the PI3K/Akt signalling pathway. *Cell Microbiol,* 11**,** 863-71.

EHRHARDT, C., MARJUKI, H., WOLFF, T., NURNBERG, B., PLANZ, O., PLESCHKA, S. & LUDWIG, S. 2006. Bivalent role of the phosphatidylinositol-3-kinase (PI3K) during influenza virus infection and host cell defence. Cell Microbiol, 8, 1336-48.

EHRHARDT, C., WOLFF, T., PLESCHKA, S., PLANZ, O., BEERMANN, W., BODE, J. G., SCHMOLKE, M. & LUDWIG, S. 2007. Influenza A virus NS1 protein activates the PI3K/Akt pathway to mediate antiapoptotic signaling responses. J Virol, 81, 3058-67.

EISFELD, A. J., NEUMANN, G. & KAWAOKA, Y. 2015. At the centre: influenza A virus ribonucleoproteins. *Nat Rev Microbiol,* 13, 28-41.

ELDERFIELD, R. A., WATSON, S. J., GODLEE, A., ADAMSON, W. E., THOMPSON, C. I., DUNNING, J., FERNANDEZ-ALONSO, M., BLUMENKRANTZ, D., HUSSELL, T., INVESTIGATORS, M., ZAMBON, M., OPENSHAW, P., KELLAM, P. & BARCLAY, W. S. 2014. Accumulation of human-adapting mutations during circulation of A(H1N1)pdm09 influenza virus in humans in the United Kingdom. J Virol, 88, 13269-83.

ELLEMAN, C. J. & BARCLAY, W. S. 2004. The M1 matrix protein controls the filamentous phenotype of influenza A virus. *Virology,* 321, 144-53.

ELLIS, M. J. & GOODBOURN, S. 1994. NF-kappa B-independent activation of beta-interferon expression in mouse F9 embryonal carcinoma cells. *Nucleic Acids Res,* 22, 4489-96.

ELTON, D. & BRYANT, N. 2011. Facing the threat of equine influenza. *Equine Vet J,* 43, 250-8.

ELTON, D., BRUCE, E. A., BRYANT, N., WISE, H. M., MACRAE, S., RASH, A., SMITH, N., TURNBULL, M. L., MEDCALF, L., DALY, J. M. & DIGARD, P. 2013. The genetics of virus particle shape in equine influenza A virus. *Influenza Other Respir Viruses,* 7 Suppl 4, 81-9.

ENAMI, K., SATO, T. A., NAKADA, S. & ENAMI, M. 1994. Influenza virus NS1 protein stimulates translation of the M1 protein. J Virol, 68, 1432-7.

ENGELMAN, J. A., LUO, J. & CANTLEY, L. C. 2006. The evolution of phosphatidylinositol 3-kinases as regulators of growth and metabolism. *Nat Rev Genet,* 7, 606-19.

ERLANDSSON, L., BLUMENTHAL, R., ELORANTA, M. L., ENGEL, H., ALM, G., WEISS, S. & LEANDERSON, T. 1998. Interferon-beta is required for interferon-alpha production in mouse fibroblasts. *Curr Biol,* 8, 223-6.

EVANS, D. L., CONNOR, M. A., MOSS, L. D., LACKAY, S., LEARY, J. H., 3RD, KRUNKOSKY, T. & JASO-FRIEDMANN, L. 2009. Cellular expression and antimicrobial function of a phylogenetically conserved novel histone 1x-like protein on mouse cells: a potential new class of pattern recognition receptor. J Leukoc Biol, 86, 133-41.

EVERETT, R. D., BOUTELL, C. & HALE, B. G. 2013. Interplay between viruses and host sumoylation pathways. *Nat Rev Microbiol,* 11, 400-11.

FALCON, A. M., FORTES, P., MARION, R. M., BELOSO, A. & ORTIN, J. 1999. Interaction of influenza virus NS1 protein and the human homologue of Staufen in vivo and in vitro. *Nucleic Acids Res,* 27, 2241-7.

FAN, H., ZHAO, G., REN, D., LIU, F., DONG, G. & HOU, Y. 2017. Gender differences of B cell signature related to estrogen-induced IFI44L/BAFF in systemic lupus erythematosus. Immunol Lett, 181, 71-78.

FENG, K. H., SUN, M., IKETANI, S., HOLMES, E. C. & PARRISH, C. R. 2016. Comparing the functions of equine and canine influenza H3N8 virus PA-X proteins: Suppression of reporter gene expression and modulation of global host gene expression. *Virology,* 496, 138-146.

FENINI, G., CONTASSOT, E. & FRENCH, L. E. 2017. Potential of IL-1, IL-18 and Inflammasome Inhibition for the Treatment of Inflammatory Skin Diseases. Front Pharmacol, 8, 278.

FINKELSHTEIN, D., WERMAN, A., NOVICK, D., BARAK, S. & RUBINSTEIN, M. 2013. LDL receptor and its family members serve as the cellular receptors for vesicular stomatitis virus. Proc Natl Acad Sci U S A, 110, 7306-11.

FOLSOM, R. W., LITTLEFIELD-CHABAUD, M. A., FRENCH, D. D., POURCIAU, S. S., MISTRIC, L. & HOROHOV, D. W. 2001. Exercise alters the immune response to equine influenza virus and increases susceptibility to infection. *Equine Vet J,* 33**,** 664-9.

FORTES, P., BELOSO, A. & ORTIN, J. 1994. Influenza virus NS1 protein inhibits pre-mRNA splicing and blocks mRNA nucleocytoplasmic transport. *EMBO J,* 13**,** 704-12.

FRAHM, T., HAUSER, H. & KOSTER, M. 2006. IFN-type-I-mediated signaling is regulated by modulation of STAT2 nuclear export. *J Cell Sci,* 119**,** 1092-104.

FRANKE, T. F., HORNIK, C. P., SEGEV, L., SHOSTAK, G. A. & SUGIMOTO, C. 2003. PI3K/Akt and apoptosis: size matters. Oncogene, 22, 8983-98.

FREZZETTI, D., GALLO, M., ROMA, C., D'ALESSIO, A., MAIELLO, M. R., BEVILACQUA, S., NORMANNO, N. & DE LUCA, A. 2016. Vascular Endothelial Growth Factor A Regulates the Secretion of Different Angiogenic Factors in Lung Cancer Cells. J Cell Physiol, 231, 1514-21.
FUJIMOTO, I., TAKIZAWA, T., OHBA, Y. & NAKANISHI, Y. 1998. Co-expression of Fas and Fas-ligand on the surface of influenza virus-infected cells. Cell Death Differ, 5, 426-31.

FUMAGALLI, F., NOACK, J., BERGMANN, T. J., CEBOLLERO, E., PISONI, G. B., FASANA, E., FREGNO, I., GALLI, C., LOI, M., SOLDA, T., D'ANTUONO, R., RAIMONDI, A., JUNG, M., MELNYK, A., SCHORR, S., SCHREIBER, A., SIMONELLI, L., VARANI, L., WILSON-ZBINDEN, C., ZERBE, O., HOFMANN, K., PETER, M., QUADRONI, M., ZIMMERMANN, R. & MOLINARI, M. 2016. Translocon component Sec62 acts in endoplasmic reticulum turnover during stress recovery. Nat Cell Biol, 18, 1173-1184.

GACK, M. U., ALBRECHT, R. A., URANO, T., INN, K. S., HUANG, I. C., CARNERO, E., FARZAN, M., INOUE, S., JUNG, J. U. & GARCIA-SASTRE, A. 2009. Influenza A virus NS1 targets the ubiquitin ligase TRIM25 to evade recognition by the host viral RNA sensor RIG-I. Cell Host Microbe, 5, 439-49.

GAGNON, C. A., ELAHI, S. M., TREMBLAY, D., LAVOIE, J. P., BRYANT, N. A., ELTON, D. M., CARMAN, S. & ELSENER, J. 2007. Genetic relatedness of recent Canadian equine influenza virus isolates with vaccine strains used in the field. Can Vet J, 48, 1028-30.

GAMSJAEGER, R., LIEW, C. K., LOUGHLIN, F. E., CROSSLEY, M. & MACKAY, J. P. 2007. Sticky fingers: zinc-fingers as protein-recognition motifs. Trends Biochem Sci, 32, 63-70.

GARAIGORTA, U. & ORTIN, J. 2007. Mutation analysis of a recombinant NS replicon shows that influenza virus NS1 protein blocks the splicing and nucleo-cytoplasmic transport of its own viral mRNA. *Nucleic Acids Res,* 35**,** 4573-82.

GARCIA-SASTRE, A., EGOROV, A., MATASSOV, D., BRANDT, S., LEVY, D. E., DURBIN, J. E., PALESE, P. & MUSTER, T. 1998. Influenza A virus lacking the NS1 gene replicates in interferon-deficient systems. Virology, 252, 324-30.

GARCIA, M. A., GIL, J., VENTOSO, I., GUERRA, S., DOMINGO, E., RIVAS, C. & ESTEBAN, M. 2006. Impact of protein kinase PKR in cell biology: from antiviral to antiproliferative action. *Microbiol Mol Biol Rev,* 70**,** 1032-60.

GARFINKEL, M. S. & KATZE, M. G. 1993. Translational control by influenza virus. Selective translation is mediated by sequences within the viral mRNA 5'-untranslated region. J Biol Chem, 268, 22223-6.

GEISS, G. K., SALVATORE, M., TUMPEY, T. M., CARTER, V. S., WANG, X., BASLER, C. F., TAUBENBERGER, J. K., BUMGARNER, R. E., PALESE, P., KATZE, M. G. & GARCIA-SASTRE, A. 2002.

Cellular transcriptional profiling in influenza A virus-infected lung epithelial cells: the role of the nonstructural NS1 protein in the evasion of the host innate defense and its potential contribution to pandemic influenza. Proc Natl Acad Sci U S A, 99, 10736-41.

GEORGE, L., MITRA, A., THIMRAJ, T. A., IRMLER, M., VISHWESWARAIAH, S., LUNDING, L., HUHN, D., MADURGA, A., BECKERS, J., FEHRENBACH, H., UPADHYAY, S., SCHULZ, H., LEIKAUF, G. D. & GANGULY, K. 2017. Transcriptomic analysis comparing mouse strains with extreme total lung capacities identifies novel candidate genes for pulmonary function. Respir Res, 18, 152.

GIBSON, D. G., BELL, S. P. & APARICIO, O. M. 2006. Cell cycle execution point analysis of ORC function and characterization of the checkpoint response to ORC inactivation in Saccharomyces cerevisiae. Genes Cells, 11, 557-73.

GIESE, S., BOLTE, H. & SCHWEMMLE, M. 2016. The Feat of Packaging Eight Unique Genome Segments. *Viruses,* 8. Pii: E165.

GIL, J. & ESTEBAN, M. 2000. Induction of apoptosis by the dsRNA-dependent protein kinase (PKR): mechanism of action. Apoptosis, 5, 107-14.

GILDEA, S., FITZPATRICK, D. A. & CULLINANE, A. 2013. Epidemiological and virological investigations of equine influenza outbreaks in Ireland (2010-2012). Influenza Other Respir Viruses, 7 Suppl 4, 61-72.

GILDEA, S., QUINLIVAN, M., ARKINS, S. & CULLINANE, A. 2012. The molecular epidemiology of equine influenza in Ireland from 2007-2010 and its international significance. *Equine Vet J,* 44**,** 387-92.

GOLD, E. S., DIERCKS, A. H., PODOLSKY, I., PODYMINOGIN, R. L., ASKOVICH, P. S., TREUTING, P. M. & ADEREM, A. 2014. 25-Hydroxycholesterol acts as an amplifier of inflammatory signaling. Proc Natl Acad Sci U S A, 111, 10666-71.

GOLEBIEWSKI, L., LIU, H., JAVIER, R. T. & RICE, A. P. 2011. The avian influenza virus NS1 ESEV PDZ binding motif associates with Dlg1 and Scribble to disrupt cellular tight junctions. *J Virol,* 85**,** 10639-48.

GONZALEZ, G., MARSHALL, J. F., MORRELL, J., ROBB, D., MCCAULEY, J. W., PEREZ, D. R., PARRISH, C. R. & MURCIA, P. R. 2014. Infection and pathogenesis of canine, equine, and human influenza viruses in canine tracheas. *Journal of virology,* 88**,** 9208-19.

GOODBOURN, S., ZINN, K. & MANIATIS, T. 1985. Human beta-interferon gene expression is regulated by an inducible enhancer element. *Cell,* 41**,** 509-20.

GOODMAN, A. G., SMITH, J. A., BALACHANDRAN, S., PERWITASARI, O., PROLL, S. C., THOMAS, M. J., KORTH, M. J., BARBER, G. N., SCHIFF, L. A. & KATZE, M. G. 2007. The cellular protein P58IPK regulates influenza virus mRNA translation and replication through a PKR-mediated mechanism. *J Virol,* 81**,** 2221-30.

GREENSPAN, D., PALESE, P. & KRYSTAL, M. 1988. Two nuclear location signals in the influenza virus NS1 nonstructural protein. *J Virol,* 62**,** 3020-6.

GRIMM, D., STAEHELI, P., HUFBAUER, M., KOERNER, I., MARTINEZ-SOBRIDO, L., SOLORZANO, A., GARCIA-SASTRE, A., HALLER, O. & KOCHS, G. 2007. Replication fitness determines high virulence of influenza A virus in mice carrying functional Mx1 resistance gene. *Proc Natl Acad Sci U S A,* 104**,** 6806-11.

GU, W., GALLAGHER, G. R., DAI, W., LIU, P., LI, R., TROMBLY, M. I., GAMMON, D. B., MELLO, C. C., WANG, J. P. & FINBERG, R. W. 2015. Influenza A virus preferentially snatches noncoding RNA caps. *RNA,* 21**,** 2067-75.

GUO, Y., WANG, M., KAWAOKA, Y., GORMAN, O., ITO, T., SAITO, T. & WEBSTER, R. G. 1992. Characterization of a new avian-like influenza A virus from horses in China. *Virology,* 188**,** 245-55.
GUO, Z., CHEN, L. M., ZENG, H., GOMEZ, J. A., PLOWDEN, J., FUJITA, T., KATZ, J. M., DONIS, R. O. & SAMBHARA, S. 2007. NS1 protein of influenza A virus inhibits the function of intracytoplasmic pathogen sensor, RIG-I. *Am J Respir Cell Mol Biol,* 36**,** 263-9.

GUO, J., WANG, W., YU, D., WU,Y. 2011. Spinoculation Triggers Dynamic Actin and Cofilin Activity That Facilitates HIV-1 Infection of Transformed and Resting CD4 T Cells. J Virol 85, 9824-9833.

HACKER, H., REDECKE, V., BLAGOEV, B., KRATCHMAROVA, I., HSU, L. C., WANG, G. G., KAMPS, M. P., RAZ, E., WAGNER, H., HACKER, G., MANN, M. & KARIN, M. 2006. Specificity in Toll-like receptor signalling through distinct effector functions of TRAF3 and TRAF6. *Nature,* 439**,** 204-7.
HALE, B. G., BARCLAY, W. S., RANDALL, R. E. & RUSSELL, R. J. 2008. Structure of an avian influenza A virus NS1 protein effector domain. Virology, 378, 1-5.

HALE, B. G., BATTY, I. H., DOWNES, C. P. & RANDALL, R. E. 2008b. Binding of influenza A virus NS1 protein to the inter-SH2 domain of p85 suggests a novel mechanism for phosphoinositide 3-kinase activation. J Biol Chem, 283, 1372-80.

HALE, B. G., JACKSON, D., CHEN, Y. H., LAMB, R. A. & RANDALL, R. E. 2006. Influenza A virus NS1 protein binds p85beta and activates phosphatidylinositol-3-kinase signaling. *Proc Natl Acad Sci U S A,* 103**,** 14194-9.

HALE, B. G., KNEBEL, A., BOTTING, C. H., GALLOWAY, C. S., PRECIOUS, B. L., JACKSON, D., ELLIOTT, R. M. & RANDALL, R. E. 2009. CDK/ERK-mediated phosphorylation of the human influenza A virus NS1 protein at threonine-215. *Virology,* 383**,** 6-11.

HALE, B. G., RANDALL, R. E., ORTIN, J. & JACKSON, D. 2008. The multifunctional NS1 protein of influenza A viruses. J Gen Virol, 89, 2359-76.

HALE, B. G., STEEL, J., MEDINA, R. A., MANICASSAMY, B., YE, J., HICKMAN, D., HAI, R., SCHMOLKE, M., LOWEN, A. C., PEREZ, D. R. & GARCIA-SASTRE, A. 2010b. Inefficient control of host gene expression by the 2009 pandemic H1N1 influenza A virus NS1 protein. J Virol, 84, 6909-22.

HALLEN, L. C., BURKI, Y., EBELING, M., BROGER, C., SIEGRIST, F., OROSZLAN-SZOVIK, K., BOHRMANN, B., CERTA, U. & FOSER, S. 2007. Antiproliferative activity of the human IFN-alpha-inducible protein IFI44. J Interferon Cytokine Res, 27, 675-80.

HALLER, O., STAEHELI, P. & KOCHS, G. 2007. Interferon-induced Mx proteins in antiviral host defense. Biochimie, 89, 812-8.

HARDY, M. P., MC, G. A. F. & O'NEILL, L. A. 2004. Transcriptional regulation of the human TRIF (TIR domain-containing adaptor protein inducing interferon beta) gene. *Biochem J,* 380**,** 83-93.

HARRIS, A., CARDONE, G., WINKLER, D. C., HEYMANN, J. B., BRECHER, M., WHITE, J. M. & STEVEN, A. C. 2006. Influenza virus pleiomorphy characterized by cryoelectron tomography. *Proc Natl Acad Sci U S A,* 103**,** 19123-7.

HARRIS, B. Z. & LIM, W. A. 2001. Mechanism and role of PDZ domains in signaling complex assembly. J Cell Sci, 114, 3219-31.

HARTL, F. U. & HAYER-HARTL, M. 2002. Molecular chaperones in the cytosol: from nascent chain to folded protein. Science, 295, 1852-8.

HATADA, E. & FUKUDA, R. 1992. Binding of influenza A virus NS1 protein to dsRNA in vitro. *J Gen Virol,* 73**,** 3325-9.

HATADA, E., SAITO, S., OKISHIO,N., FUKUDA, R. 1997. Binding of the influenza virus NS1 protein to model genome RNAs. *J Gen Virol,* 78**,** 1059-1063.

HATADA, E., SAITO, S. & FUKUDA, R. 1999. Mutant influenza viruses with a defective NS1 protein cannot block the activation of PKR in infected cells. J Virol, 73, 2425-33.

HAYMAN, A., COMELY, S., LACKENBY, A., HARTGROVES, L. C., GOODBOURN, S., MCCAULEY, J. W. & BARCLAY, W. S. 2007. NS1 proteins of avian influenza A viruses can act as antagonists of the human alpha/beta interferon response. J Virol, 81, 2318-27.

HAYMAN, A., COMELY, S., LACKENBY, A., MURPHY, S., McCauley, J., GOODBOURN, S., & BARCLAY, W. S. 2006. Variation in the ability of human influenza A viruses to induce and inhibit the IFN-beta pathway. Virology, 347, 52-64.

HEIM, A., RYMARCZYK, B. & MAYER, T. U. 2017. Regulation of Cell Division. Adv Exp Med Biol, 953, 83-116.

HERMANNS, H. M. 2015. Oncostatin M and interleukin-31: Cytokines, receptors, signal transduction and physiology. Cytokine Growth Factor Rev, 26, 545-58.

HERREN, T., SWAISGOOD, C. & PLOW, E. F. 2003. Regulation of plasminogen receptors. Front Biosci, 8, d1-8.

HOELZER, K., MURCIA, P. R., BAILLIE, G. J., WOOD, J. L., METZGER, S. M., OSTERRIEDER, N., DUBOVI, E. J., HOLMES, E. C. & PARRISH, C. R. 2010. Intrahost evolutionary dynamics of canine influenza virus in naive and partially immune dogs. *J Virol,* 84, 5329-35.

HOFFMANN, E., STECH, J., GUAN, Y., WEBSTER, R. G. & PEREZ, D. R. 2001. Universal primer set for the full-length amplification of all influenza A viruses. Archives of virology, 146, 2275-89.

HORNUNG, V., ELLEGAST, J., KIM, S., BRZOZKA, K., JUNG, A., KATO, H., POECK, H., AKIRA, S., CONZELMANN, K. K., SCHLEE, M., ENDRES, S. & HARTMANN, G. 2006. 5'-Triphosphate RNA is the ligand for RIG-I. *Science,* 314, 994-7.

HOSSAIN, M. J., HICKMAN, D. & PEREZ, D. R. 2008. Evidence of expanded host range and mammalian-associated genetic changes in a duck H9N2 influenza virus following adaptation in quail and chickens. PLoS One, 3, e3170.

HOVANESSIAN, A. G. 1989. The double stranded RNA-activated protein kinase induced by interferon: dsRNA-PK. *J Interferon Res,* 9, 641-7.

HSIANG, T. Y., ZHOU, L. & KRUG, R. M. 2012. Roles of the phosphorylation of specific serines and threonines in the NS1 protein of human influenza A viruses. *J Virol,* 86, 10370-6.

HUANG, K. H., WANG, C. H., LEE, K. Y., LIN, S. M., LIN, C. H. & KUO, H. P. 2013. NF-kappaB repressing factor inhibits chemokine synthesis by peripheral blood mononuclear cells and alveolar macrophages in active pulmonary tuberculosis. PLoS One, 8, e77789.

HUGHES, J., ALLEN, R. C., BAGUELIN, M., HAMPSON, K., BAILLIE, G. J., ELTON, D., NEWTON, J. R., KELLAM, P., WOOD, J. L., HOLMES, E. C. & MURCIA, P. R. 2012. Transmission of equine influenza virus during an outbreak is characterized by frequent mixed infections and loose transmission bottlenecks. *PLoS Pathog,* 8, e1003081.

HUH, J. W., KIM, D. S., HA, H. S., LEE, J. R., KIM, Y. J., AHN, K., LEE, S. R., CHANG, K. T. & KIM, H. S. 2008. Cooperative exonization of MaLR and AluJo elements contributed an alternative promoter and novel splice variants of RNF19. Gene, 424, 63-70.

HUNG, A. Y. & SHENG, M. 2002. PDZ domains: structural modules for protein complex assembly. J Biol Chem, 277, 5699-702.

HUSSER, L., ALVES, M. P., RUGGLI, N. & SUMMERFIELD, A. 2011. Identification of the role of RIG-I, MDA-5 and TLR3 in sensing RNA viruses in porcine epithelial cells using lentivirus-driven RNA interference. *Virus Res,* 159, 9-16.

HUTCHINSON, E. C., CHARLES, P. D., HESTER, S. S., THOMAS, B., TRUDGIAN, D., MARTINEZ-ALONSO, M. & FODOR, E. 2014. Conserved and host-specific features of influenza virion architecture. Nat Commun, 5, 4816.

HUTCHINSON, E. C., DENHAM, E. M., THOMAS, B., TRUDGIAN, D. C., HESTER, S. S., RIDLOVA, G., YORK, A., TURRELL, L. & FODOR, E. 2012. Mapping the phosphoproteome of influenza A and B viruses by mass spectrometry. *PLoS Pathog,* 8**,** e1002993.

HUTCHINSON, E.C., FODOR, E. 2014. Purification of influenza virions by haemadsorption and ultracentrifugation. Protocol Exchange.

HUTCHINSON, E. C., VON KIRCHBACH, J. C., GOG, J. R. & DIGARD, P. 2010. Genome packaging in influenza A virus. *J Gen Virol,* 91**,** 313-28.

HYLAND, L., WEBBY, R., SANDBULTE, M. R., CLARKE, B. & HOU, S. 2006. Influenza virus NS1 protein protects against lymphohematopoietic pathogenesis in an in vivo mouse model. *Virology,* 349**,** 156-63.

ICHIHARA-TANAKA, K., KADOMATSU, K. & KISHIDA, S. 2017. Temporally and Spatially Regulated Expression of the Linker Histone H1fx During Mouse Development. J Histochem Cytochem, 65, 513-530.

IMAI, M. & KAWAOKA, Y. 2012. The role of receptor binding specificity in interspecies transmission of influenza viruses. *Curr Opin Virol,* 2**,** 160-7.

INGLIS, S. C., BARRETT, T., BROWN, C. M. & ALMOND, J. W. 1979. The smallest genome RNA segment of influenza virus contains two genes that may overlap. *Proc Natl Acad Sci U S A,* 76**,** 3790-4.

IOANNIDIS, I., YE, F., MCNALLY, B., WILLETTE, M. & FLANO, E. 2013. Toll-like receptor expression and induction of type I and type III interferons in primary airway epithelial cells. *J Virol,* 87**,** 3261-70.

ISAACS, A. & LINDENMANN, J. 1957. Virus interference. I. The interferon. *Proc R Soc Lond B Biol Sci,* 147**,** 258-67.

ITO, M., NAGAI, M., HAYAKAWA, Y., KOMAE, H., MURAKAMI, N., YOTSUYA, S., ASAKURA, S., SAKODA, Y. & KIDA, H. 2008. Genetic Analyses of an H3N8 Influenza Virus Isolate, Causative Strain of the Outbreak of Equine Influenza at the Kanazawa Racecourse in Japan in 2007. *J Vet Med Sci,* 70**,** 899-906.

IWAI, A., SHIOZAKI, T., KAWAI, T., AKIRA, S., KAWAOKA, Y., TAKADA, A., KIDA, H. & MIYAZAKI, T. 2010. Influenza A virus polymerase inhibits type I interferon induction by binding to interferon beta promoter stimulator 1. *J Biol Chem,* 285**,** 32064-74.

IWASAKI, A. & PILLAI, P. S. 2014. Innate immunity to influenza virus infection. *Nat Rev Immunol,* 14**,** 315-28.

JACKSON, D., HOSSAIN, M. J., HICKMAN, D., PEREZ, D. R. & LAMB, R. A. 2008. A new influenza virus virulence determinant: the NS1 protein four C-terminal residues modulate pathogenicity. *Proc Natl Acad Sci U S A,* 105**,** 4381-6.

JAGGER, B. W., WISE, H. M., KASH, J. C., WALTERS, K. A., WILLS, N. M., XIAO, Y. L., DUNFEE, R. L., SCHWARTZMAN, L. M., OZINSKY, A., BELL, G. L., DALTON, R. M., LO, A., EFSTATHIOU, S., ATKINS, J. F., FIRTH, A. E., TAUBENBERGER, J. K. & DIGARD, P. 2012. An overlapping protein-coding region in influenza A virus segment 3 modulates the host response. *Science,* 337**,** 199-204.

JANG, L. K., LEE, Z. H., KIM, H. H., HILL, J. M., KIM, J. D. & KWON, B. S. 2001. A novel leucine-rich repeat protein (LRR-1): potential involvement in 4-1BB-mediated signal transduction. Mol Cells, 12, 304-12.

JAVIER, R. T. & RICE, A. P. 2011. Emerging theme: cellular PDZ proteins as common targets of pathogenic viruses. J Virol, 85, 11544-56.

JEISY-SCOTT, V., KIM, J. H., DAVIS, W. G., CAO, W., KATZ, J. M. & SAMBHARA, S. 2012. TLR7 recognition is dispensable for influenza virus A infection but important for the induction of hemagglutinin-specific antibodies in response to the 2009 pandemic split vaccine in mice. *J Virol,* 86**,** 10988-98.

JIAO, P., TIAN, G., LI, Y., DENG, G., JIANG, Y., LIU, C., LIU, W., BU, Z., KAWAOKA, Y. & CHEN, H. 2008. A single-amino-acid substitution in the NS1 protein changes the pathogenicity of H5N1 avian influenza viruses in mice. J Virol, 82, 1146-54.

JIN, J., CARDOZO, T., LOVERING, R. C., ELLEDGE, S. J., PAGANO, M. & HARPER, J. W. 2004. Systematic analysis and nomenclature of mammalian F-box proteins. Genes Dev, 18, 2573-80.

JONGES, M., WELKERS, M. R., JEENINGA, R. E., MEIJER, A., SCHNEEBERGER, P., FOUCHIER, R. A., DE JONG, M. D. & KOOPMANS, M. 2014. Emergence of the virulence-associated PB2 E627K substitution in a fatal human case of highly pathogenic avian influenza virus A(H7N7) infection as determined by Illumina ultra-deep sequencing. *J Virol,* 88**,** 1694-702.

JOSSET, L., ZENG, H., KELLY, S. M., TUMPEY, T. M. & KATZE, M. G. 2014. Transcriptomic characterization of the novel avian-origin influenza A (H7N9) virus: specific host response and responses intermediate between avian (H5N1 and H7N7) and human (H3N2) viruses and implications for treatment options. MBio, 5, e01102-13.

KAMBACH, C., WALKE, S., YOUNG, R., AVIS, J. M., DE LA FORTELLE, E., RAKER, V. A., LUHRMANN, R., LI, J. & NAGAI, K. 1999. Crystal structures of two Sm protein complexes and their implications for the assembly of the spliceosomal snRNPs. Cell, 96, 375-87.

KAMINSKI, D. A., WEI, C., QIAN, Y., ROSENBERG, A. F. & SANZ, I. 2012. Advances in human B cell phenotypic profiling. *Front Immunol,* 3**,** 302.

KARAMENDIN, K., KYDYRMANOV, A., SAYATOV, M., STROCHKOV, V., SANDYBAYEV, N. & SULTANKULOVA, K. 2016. Retrospective Analysis of the Equine Influenza Virus A/Equine/Kirgizia/26/1974 (H7N7) Isolated in Central Asia. *Pathogens,* 5.

KASEL, J. A., ALFORD, R. H., KNIGHT, V., WADDELL, G. H. & SIGEL, M. M. 1965. Experimental Infection of Human Volunteers with Equine Influenza Virus. *Nature,* 206**,** 41-3.

KATO, H., TAKEUCHI, O., SATO, S., YONEYAMA, M., YAMAMOTO, M., MATSUI, K., UEMATSU, S., JUNG, A., KAWAI, T., ISHII, K. J., YAMAGUCHI, O., OTSU, K., TSUJIMURA, T., KOH, C. S., REIS E SOUSA, C., MATSUURA, Y., FUJITA, T. & AKIRA, S. 2006. Differential roles of MDA5 and RIG-I helicases in the recognition of RNA viruses. Nature, 441, 101-5.

KATZE, M. G. & KRUG, R. M. 1984. Metabolism and expression of RNA polymerase II transcripts in influenza virus-infected cells. Mol Cell Biol, 4, 2198-206.

KATZE, M. G., DETJEN, B. M., SAFER, B. & KRUG, R. M. 1986. Translational control by influenza virus: suppression of the kinase that phosphorylates the alpha subunit of initiation factor eIF-2 and selective translation of influenza viral mRNAs. *Mol Cell Biol,* 6**,** 1741-50.

KATZE, M. G., TOMITA, J., BLACK, T., KRUG, R. M., SAFER, B. & HOVANESSIAN, A. 1988. Influenza virus regulates protein synthesis during infection by repressing autophosphorylation and activity of the cellular 68,000-Mr protein kinase. *J Virol,* 62**,** 3710-7.

KAWAI, T., TAKAHASHI, K., SATO, S., COBAN, C., KUMAR, H., KATO, H., ISHII, K. J., TAKEUCHI, O. & AKIRA, S. 2005. IPS-1, an adaptor triggering RIG-I- and Mda5-mediated type I interferon induction. *Nat Immunol,* 6, 981-8.

KAWAOKA, Y., BEAN, W. J. & WEBSTER, R. G. 1989. Evolution of the hemagglutinin of equine H3 influenza viruses. *Virology,* 169, 283-92.

KAWAOKA, Y., GORMAN, O. T., ITO, T., WELLS, K., DONIS, R. O., CASTRUCCI, M. R., DONATELLI, I. & WEBSTER, R. G. 1998. Influence of host species on the evolution of the nonstructural (NS) gene of influenza A viruses. *Virus Res,* 55, 143-56.

KECHAGIA, M., PAPASSOTIRIOU, I. & GOURGOULIANIS, K. I. 2016. Endocan and the respiratory system: a review. Int J Chron Obstruct Pulmon Dis, 11, 3179-3187.

KERRY, P. S., AYLLON, J., TAYLOR, M. A., HASS, C., LEWIS, A., GARCIA-SASTRE, A., RANDALL, R. E., HALE, B. G. & RUSSELL, R. J. 2011. A transient homotypic interaction model for the influenza A virus NS1 protein effector domain. PLoS One, 6, e17946.
KERRY, P. S., LONG, E., TAYLOR, M. A. & RUSSELL, R. J. 2011b. Conservation of a crystallographic interface suggests a role for beta-sheet augmentation in influenza virus NS1 multifunctionality. *Acta Crystallogr Sect F Struct Biol Cryst Commun,* 67, 858-61.

KHSHEIBUN, R., PAPERNA, T., VOLKOWICH, A., LEJBKOWICZ, I., AVIDAN, N. & MILLER, A. 2014. Gene expression profiling of the response to interferon beta in Epstein-Barr-transformed and primary B cells of patients with multiple sclerosis. PLoS One, 9, e102331.

KHURELBAATAR, N., KRUEGER, W. S., HEIL, G. L., DARMAA, B., ULZIIMAA, D., TSERENNOROV, D., BATERDENE, A., ANDERSON, B. D. & GRAY, G. C. 2014. Little evidence of avian or equine influenza virus infection among a cohort of Mongolian adults with animal exposures, 2010-2011. *PLoS One,* 9, e85616.

KILLIP, M. J., FODOR, E. & RANDALL, R. E. 2015. Influenza virus activation of the interferon system. *Virus Res,* 209, 11-22.

KIM, E. & SHENG, M. 2004. PDZ domain proteins of synapses. Nat Rev Neurosci, 5, 771-81.

KIM, J. I., HWANG, M. W., LEE, I., PARK, S., LEE, S., BAE, J. Y., HEO, J., KIM, D., JANG, S. I., PARK, M. S., KWON, H. J., SONG, J. W. & PARK, M. S. 2014. The PDZ-binding motif of the avian NS1 protein affects transmission of the 2009 influenza A(H1N1) virus. Biochem Biophys Res Commun, 449, 19-25.

KING, P. & GOODBOURN, S. 1994. The beta-interferon promoter responds to priming through multiple independent regulatory elements. *J Biol Chem,* 269, 30609-15.

KIRKLAND, P. D., FINLAISON, D. S., CRISPE, E. & HURT, A. C. 2010. Influenza virus transmission from horses to dogs, Australia. *Emerg Infect Dis,* 16, 699-702.

KOCHS, G., GARCIA-SASTRE, A. & MARTINEZ-SOBRIDO, L. 2007. Multiple anti-interferon actions of the influenza A virus NS1 protein. J Virol, 81, 7011-21.

KOCHS, G., KOERNER, I., THIEL, L., KOTHLOW, S., KASPERS, B., RUGGLI, N., SUMMERFIELD, A., PAVLOVIC, J., STECH, J. & STAEHELI, P. 2007c. Properties of H7N7 influenza A virus strain SC35M lacking interferon antagonist NS1 in mice and chickens. *J Gen Virol,* 88, 1403-9.

KOPPSTEIN, D., ASHOUR, J. & BARTEL, D. P. 2015. Sequencing the cap-snatching repertoire of H1N1 influenza provides insight into the mechanism of viral transcription initiation. *Nucleic Acids Res,* 43, 5052-64.

KOWALINSKI, E., LUNARDI, T., MCCARTHY, A. A., LOUBER, J., BRUNEL, J., GRIGOROV, B., GERLIER, D. & CUSACK, S. 2011. Structural basis for the activation of innate immune pattern-recognition receptor RIG-I by viral RNA. *Cell,* 147, 423-35.

KRONCKE, K. D. & KLOTZ, L. O. 2009. Zinc fingers as biologic redox switches? Antioxid Redox Signal, 11, 1015-27.

KRUG, R. M. & ETKIND, P. R. 1973. Cytoplasmic and nuclear virus-specific proteins in influenza virus-infected MDCK cells. *Virology,* 56, 334-48.

KRUG, R. M. 2015. Functions of the influenza A virus NS1 protein in antiviral defense. Curr Opin Virol, 12, 1-6.

KUBALLA, P., BAUMANN, A. L., MAYER, K., BAR, U., BURTSCHER, H. & BRINKMANN, U. 2015. Induction of heat shock protein HSPA6 (HSP70B') upon HSP90 inhibition in cancer cell lines. FEBS Lett, 589, 1450-8.

KUNITA, R., OTOMO, A. & IKEDA, J. E. 2002. Identification and characterization of novel members of the CREG family, putative secreted glycoproteins expressed specifically in brain. Genomics, 80, 456-60.

KUO, R. L., ZHAO, C., MALUR, M. & KRUG, R. M. 2010. Influenza A virus strains that circulate in humans differ in the ability of their NS1 proteins to block the activation of IRF3 and interferon-beta transcription. Virology, 408, 146-58.

KURE, S., KOJIMA, K., KUDO, T., KANNO, K., AOKI, Y., SUZUKI, Y., SHINKA, T., SAKATA, Y., NARISAWA, K. & MATSUBARA, Y. 2001. Chromosomal localization, structure, single-nucleotide polymorphisms, and expression of the human H-protein gene of the glycine cleavage system (GCSH), a candidate gene for nonketotic hyperglycinemia. J Hum Genet, 46, 378-84.

KUSS-DUERKOP, S. K., WANG, J., MENA, I., WHITE, K., METREVELI, G., SAKTHIVEL, R., MATA, M. A., MUNOZ-MORENO, R., CHEN, X., KRAMMER, F., DIAMOND, M. S., CHEN, Z. J., GARCIA-SASTRE, A. & FONTOURA, B. M. A. 2017. Influenza virus differentially activates mTORC1 and mTORC2 signaling to maximize late stage replication. PLoS Pathog, 13, e1006635.

LAI, A. C., CHAMBERS, T. M., HOLLAND, R. E., JR., MORLEY, P. S., HAINES, D. M., TOWNSEND, H. G. & BARRANDEGUY, M. 2001. Diverged evolution of recent equine-2 influenza (H3N8) viruses in the Western Hemisphere. *Arch Virol,* 146, 1063-74.

LAI, A. C., ROGERS, K. M., GLASER, A., TUDOR, L. & CHAMBERS, T. 2004. Alternate circulation of recent equine-2 influenza viruses (H3N8) from two distinct lineages in the United States. *Virus Res,* 100, 159-64.

LAKDAWALA, S. S., WU, Y., WAWRZUSIN, P., KABAT, J., BROADBENT, A. J., LAMIRANDE, E. W., FODOR, E., ALTAN-BONNET, N., SHROFF, H. & SUBBARAO, K. 2014. Influenza a virus assembly intermediates fuse in the cytoplasm. *PLoS Pathog,* 10, e1003971.

LAMB, R. A. & CHOPPIN, P. W. 1979. Segment 8 of the influenza virus genome is unique in coding for two polypeptides. Proc Natl Acad Sci U S A, 76, 4908-12.

LAMB, R. A. & LAI, C. J. 1980. Sequence of interrupted and uninterrupted mRNAs and cloned DNA coding for the two overlapping nonstructural proteins of influenza virus. *Cell,* 21, 475-85.

LANDOLT, G. A. 2014. Equine influenza virus. *Vet Clin North Am Equine Pract,* 30, 507-22.

LAZAROWITZ, S. G., COMPANS, R. W. & CHOPPIN, P. W. 1971. Influenza virus structural and nonstructural proteins in infected cells and their plasma membranes. *Virology,* 46, 830-43.

LE GOFFIC, R., POTHLICHET, J., VITOUR, D., FUJITA, T., MEURS, E., CHIGNARD, M. & SI-TAHAR, M. 2007. Cutting Edge: Influenza A virus activates TLR3-dependent inflammatory and RIG-I-dependent antiviral responses in human lung epithelial cells. *J Immunol,* 178, 3368-72.

LE, Q. M., SAKAI-TAGAWA, Y., OZAWA, M., ITO, M. & KAWAOKA, Y. 2009. Selection of H5N1 influenza virus PB2 during replication in humans. Journal of virology, 83, 5278-81.

LEE, C. Y., TSAI, Y. T., LOH, S. H., LIU, J. C., CHEN, T. H., CHAO, H. H., CHENG, T. H. & CHEN, J. J. 2014. Urotensin II induces interleukin 8 expression in human umbilical vein endothelial cells. PLoS One, 9, e90278.

LEE, S. T., FENG, M., WEI, Y., LI, Z., QIAO, Y., GUAN, P., JIANG, X., WONG, C. H., HUYNH, K., WANG, J., LI, J., KARUTURI, K. M., TAN, E. Y., HOON, D. S., KANG, Y. & YU, Q. 2013. Protein tyrosine phosphatase UBASH3B is overexpressed in triple-negative breast cancer and promotes invasion and metastasis. Proc Natl Acad Sci U S A, 110, 11121-6.

LEE, S. W., JUNG, K. H., JEONG, C. H., SEO, J. H., YOON, D. K., SUH, J. K., KIM, K. W. & KIM, W. J. 2011. Inhibition of endothelial cell migration through the downregulation of MMP-9 by A-kinase anchoring protein 12. Mol Med Rep, 4, 145-9.

LEGRAND, L. J., PITEL, P. H., MARCILLAUD-PITEL, C. J., CULLINANE, A. A., COUROUCE, A. M., FORTIER, G. D., FREYMUTH, F. L. & PRONOST, S. L. 2013. Surveillance of equine influenza viruses through the RESPE network in France from November 2005 to October 2010. *Equine Vet J,* 45**,** 776-83.

LENCI, R. E., RACHAKONDA, P. S., KUBARENKO, A. V., WEBER, A. N., BRANDT, A., GAST, A., SUCKER, A., HEMMINKI, K., SCHADENDORF, D. & KUMAR, R. 2012. Integrin genes and susceptibility to human melanoma. Mutagenesis, 27, 367-73.

LEWIS, N. S., DALY, J. M., RUSSELL, C. A., HORTON, D. L., SKEPNER, E., BRYANT, N. A., BURKE, D. F., RASH, A. S., WOOD, J. L., CHAMBERS, T. M., FOUCHIER, R. A., MUMFORD, J. A., ELTON, D. M. & SMITH, D. J. 2011. Antigenic and genetic evolution of equine influenza A (H3N8) virus from 1968 to 2007. *J Virol,* 85**,** 12742-9.

LI, S., MIN, J. Y., KRUG, R. M. & SEN, G. C. 2006. Binding of the influenza A virus NS1 protein to PKR mediates the inhibition of its activation by either PACT or double-stranded RNA. *Virology,* 349**,** 13-21.

LI, Y., BANERJEE, S., WANG, Y., GOLDSTEIN, S. A., DONG, B., GAUGHAN, C., SILVERMAN, R. H. & WEISS, S. R. 2016a. Activation of RNase L is dependent on OAS3 expression during infection with diverse human viruses. Proc Natl Acad Sci U S A, 113, 2241-6.

LI, Y., BASAVAPPA, M., LU, J., DONG, S., CRONKITE, D. A., PRIOR, J. T., REINECKER, H. C., HERTZOG, P., HAN, Y., LI, W. X., CHELOUFI, S., KARGINOV, F. V., DING, S. W. & JEFFREY, K. L. 2016b. Induction and suppression of antiviral RNA interference by influenza A virus in mammalian cells. Nat Microbiol, 2, 16250.

LI, Y., CHEN, Z. Y., WANG, W., BAKER, C. C. & KRUG, R. M. 2001. The 3'-end-processing factor CPSF is required for the splicing of single-intron pre-mRNAs in vivo. RNA, 7, 920-31.

LI, Y., YAMAKITA, Y. & KRUG, R. M. 1998. Regulation of a nuclear export signal by an adjacent inhibitory sequence: the effector domain of the influenza virus NS1 protein. *Proc Natl Acad Sci U S A,* 95**,** 4864-9.

LI, Z., JIANG, Y., JIAO, P., WANG, A., ZHAO, F., TIAN, G., WANG, X., YU, K., BU, Z. & CHEN, H. 2006. The NS1 gene contributes to the virulence of H5N1 avian influenza viruses. J Virol, 80, 11115-23.

LIGHTOWLERS, R. N., ROZANSKA, A. & CHRZANOWSKA-LIGHTOWLERS, Z. M. 2014. Mitochondrial protein synthesis: figuring the fundamentals, complexities and complications, of mammalian mitochondrial translation. FEBS Lett, 588, 2496-503.

LIN, C., HOLLAND, R. E., JR., DONOFRIO, J. C., MCCOY, M. H., TUDOR, L. R. & CHAMBERS, T. M. 2002. Caspase activation in equine influenza virus induced apoptotic cell death. *Vet Microbiol,* 84**,** 357-65.

LINIGER, M., SUMMERFIELD, A., ZIMMER, G., MCCULLOUGH, K. C. & RUGGLI, N. 2012. Chicken cells sense influenza A virus infection through MDA5 and CARDIF signaling involving LGP2. *J Virol,* 86**,** 705-17.

LIPSITCH, M., BARCLAY, W., RAMAN, R., RUSSELL, C. J., BELSER, J. A., COBEY, S., KASSON, P. M., LLOYD-SMITH, J. O., MAURER-STROH, S., RILEY, S., BEAUCHEMIN, C. A., BEDFORD, T., FRIEDRICH, T. C., HANDEL, A., HERFST, S., MURCIA, P. R., ROCHE, B., WILKE, C. O. & RUSSELL, C. A. 2016. Viral factors in influenza pandemic risk assessment. eLife, 5. Pii: e18491.

LIU, H., GOLEBIEWSKI, L., DOW, E. C., KRUG, R. M., JAVIER, R. T. & RICE, A. P. 2010. The ESEV PDZ-binding motif of the avian influenza A virus NS1 protein protects infected cells from apoptosis by directly targeting Scribble. *J Virol,* 84**,** 11164-74.

LIU, J. & MCFADDEN, G. 2015. SAMD9 is an innate antiviral host factor with stress response properties that can be antagonized by poxviruses. J Virol, 89, 1925-31.
LIU, J., LYNCH, P. A., CHIEN, C. Y., MONTELIONE, G. T., KRUG, R. M. & BERMAN, H. M. 1997. Crystal structure of the unique RNA-binding domain of the influenza virus NS1 protein. *Nat Struct Biol,* 4**,** 896-9.

LIU, T., CHEN, H., KIM, H., HUEN, M. S., CHEN, J. & HUANG, J. 2012. RAD18-BRCTx interaction is required for efficient repair of UV-induced DNA damage. DNA Repair (Amst), 11, 131-8.

LIU, X. Y., YANG, Z. H., PAN, X. J., ZHU, M. X. & XIE, J. P. 2010. Gene expression profile and cytotoxicity of human bronchial epithelial cells exposed to crotonaldehyde. Toxicol Lett, 197, 113-22.

LIVESAY, G. J., O'NEILL, T., HANNANT, D., YADAV, M. P. & MUMFORD, J. A. 1993. The outbreak of equine influenza (H3N8) in the United Kingdom in 1989: diagnostic use of an antigen capture ELISA. *Vet Rec,* 133**,** 515-9.

LLOYD, C. M. & MARSLAND, B. J. 2017. Lung Homeostasis: Influence of Age, Microbes, and the Immune System. Immunity, 46, 549-561.

LO, S. H. 2004. Tensin. Int J Biochem Cell Biol, 36, 31-4.

LONG, J. C. & FODOR, E. 2016. The PB2 Subunit of the Influenza A Virus RNA Polymerase Is Imported into the Mitochondrial Matrix. *J Virol,* 90**,** 8729-38.

LONG, J. S., GIOTIS, E. S., MONCORGE, O., FRISE, R., MISTRY, B., JAMES, J., MORISSON, M., IQBAL, M., VIGNAL, A., SKINNER, M. A. & BARCLAY, W. S. 2016. Species difference in ANP32A underlies influenza A virus polymerase host restriction. *Nature,* 529**,** 101-4.

LOO, Y. M., FORNEK, J., CROCHET, N., BAJWA, G., PERWITASARI, O., MARTINEZ-SOBRIDO, L., AKIRA, S., GILL, M. A., GARCIA-SASTRE, A., KATZE, M. G. & GALE, M., JR. 2008. Distinct RIG-I and MDA5 signaling by RNA viruses in innate immunity. *J Virol,* 82**,** 335-45.

LUDWIG, S., PLANZ, O., PLESCHKA, S. & WOLFF, T. 2003. Influenza-virus-induced signaling cascades: targets for antiviral therapy? Trends Mol Med, 9, 46-52.

LUDWIG, S., PLESCHKA, S., PLANZ, O. & WOLFF, T. 2006. Ringing the alarm bells: signalling and apoptosis in influenza virus infected cells. Cell Microbiol, 8, 375-86.

LUDWIG, S., SCHULTZ, U., MANDLER, J., FITCH, W. M. & SCHOLTISSEK, C. 1991. Phylogenetic relationship of the nonstructural (NS) genes of influenza A viruses. *Virology,* 183**,** 566-77.

LUDWIG, S., WANG, X., EHRHARDT, C., ZHENG, H., DONELAN, N., PLANZ, O., PLESCHKA, S., GARCIA-SASTRE, A., HEINS, G. & WOLFF, T. 2002. The influenza A virus NS1 protein inhibits activation of Jun N-terminal kinase and AP-1 transcription factors. *J Virol,* 76**,** 11166-71.

LUND, J. M., ALEXOPOULOU, L., SATO, A., KAROW, M., ADAMS, N. C., GALE, N. W., IWASAKI, A. & FLAVELL, R. A. 2004. Recognition of single-stranded RNA viruses by Toll-like receptor 7. *Proc Natl Acad Sci U S A,* 101**,** 5598-603.

LYN, R. K., SINGARAVELU, R., KARGMAN, S., O'HARA, S., CHAN, H., OBALLA, R., HUANG, Z., JONES, D. M., RIDSDALE, A., RUSSELL, R. S., PARTRIDGE, A. W. & PEZACKI, J. P. 2014. Stearoyl-CoA desaturase inhibition blocks formation of hepatitis C virus-induced specialized membranes. Sci Rep, 4, 4549.

MA, D., WU, P., EGAN, R. W., BILLAH, M. M. & WANG, P. 1999. Phosphodiesterase 4B gene transcription is activated by lipopolysaccharide and inhibited by interleukin-10 in human monocytes. Mol Pharmacol, 55, 50-7.

MAAMARY, J., PICA, N., BELICHA-VILLANUEVA, A., CHOU, Y. Y., KRAMMER, F., GAO, Q., GARCIA-SASTRE, A. & PALESE, P. 2012. Attenuated influenza virus construct with enhanced hemagglutinin protein expression. *J Virol,* 86**,** 5782-90.

MACKENZIE, K. J., CARROLL, P., MARTIN, C. A., MURINA, O., FLUTEAU, A., SIMPSON, D. J., OLOVA, N., SUTCLIFFE, H., RAINGER, J. K., LEITCH, A., OSBORN, R. T., WHEELER, A. P., NOWOTNY, M., GILBERT, N., CHANDRA, T., REIJNS, M. A. M. & JACKSON, A. P. 2017. cGAS surveillance of micronuclei links genome instability to innate immunity. Nature, 548, 461-465.

MAESTRE, A. M., GARZON, A. & RODRIGUEZ, D. 2011. Equine torovirus (BEV) induces caspase-mediated apoptosis in infected cells. PLoS One, 6, e20972.

MALATHI, K., DONG, B., GALE, M., JR. & SILVERMAN, R. H. 2007. Small self-RNA generated by RNase L amplifies antiviral innate immunity. *Nature,* 448**,** 816-9.

MALEVA KOSTOVSKA, I., WANG, J., BOGDANOVA, N., SCHURMANN, P., BHUJU, S., GEFFERS, R., DURST, M., LIEBRICH, C., KLAPDOR, R., CHRISTIANSEN, H., PARK-SIMON, T. W., HILLEMANNS, P., PLASESKA-KARANFILSKA, D. & DORK, T. 2016. Rare ATAD5 missense variants in breast and ovarian cancer patients. Cancer Lett, 376, 173-7.

MALUR, M., GALE, M., JR. & KRUG, R. M. 2012. LGP2 downregulates interferon production during infection with seasonal human influenza A viruses that activate interferon regulatory factor 3. J Virol, 86, 10733-8.

MARAZZI, I., HO, J. S., KIM, J., MANICASSAMY, B., DEWELL, S., ALBRECHT, R. A., SEIBERT, C. W., SCHAEFER, U., JEFFREY, K. L., PRINJHA, R. K., LEE, K., GARCIA-SASTRE, A., ROEDER, R. G. & TARAKHOVSKY, A. 2012. Suppression of the antiviral response by an influenza histone mimic. *Nature,* 483**,** 428-33.

MARIE, I., DURBIN, J. E. & LEVY, D. E. 1998. Differential viral induction of distinct interferon-alpha genes by positive feedback through interferon regulatory factor-7. *EMBO J,* 17**,** 6660-9.

MARION, R. M., ARAGON, T., BELOSO, A., NIETO, A. & ORTIN, J. 1997. The N-terminal half of the influenza virus NS1 protein is sufficient for nuclear retention of mRNA and enhancement of viral mRNA translation. *Nucleic Acids Res,* 25**,** 4271-7.

MARTINEZ-SOBRIDO, L., ZUNIGA, E. I., ROSARIO, D., GARCIA-SASTRE, A. & DE LA TORRE, J. C. 2006. Inhibition of the type I interferon response by the nucleoprotein of the prototypic arenavirus lymphocytic choriomeningitis virus. Journal of virology, 80, 9192-9.

MAYER, A. K., MUEHMER, M., MAGES, J., GUEINZIUS, K., HESS, C., HEEG, K., BALS, R., LANG, R. & DALPKE, A. H. 2007. Differential recognition of TLR-dependent microbial ligands in human bronchial epithelial cells. *J Immunol,* 178**,** 3134-42.

MCALEER, J. P. & KOLLS, J. K. 2014. Directing traffic: IL-17 and IL-22 coordinate pulmonary immune defense. Immunol Rev, 260, 129-44.

MCBRIDE, K. M., BANNINGER, G., MCDONALD, C. & REICH, N. C. 2002. Regulated nuclear import of the STAT1 transcription factor by direct binding of importin-alpha. *EMBO J,* 21, 1754-63.

MEDINA, R. A., STERTZ, S., MANICASSAMY, B., ZIMMERMANN, P., SUN, X., ALBRECHT, R. A., UUSI-KERTTULA, H., ZAGORDI, O., BELSHE, R. B., FREY, S. E., TUMPEY, T. M. & GARCIA-SASTRE, A. 2013. Glycosylations in the globular head of the hemagglutinin protein modulate the virulence and antigenic properties of the H1N1 influenza viruses. *Sci Transl Med,* 5, 187ra70.

MELEN, K., KINNUNEN, L., FAGERLUND, R., IKONEN, N., TWU, K. Y., KRUG, R. M. & JULKUNEN, I. 2007. Nuclear and nucleolar targeting of influenza A virus NS1 protein: striking differences between different virus subtypes. *J Virol,* 81, 5995-6006.

MELVILLE, M. W., TAN, S. L., WAMBACH, M., SONG, J., MORIMOTO, R. I. & KATZE, M. G. 1999. The cellular inhibitor of the PKR protein kinase, P58(IPK), is an influenza virus-activated co-chaperone that modulates heat shock protein 70 activity. *J Biol Chem,* 274, 3797-803.

MENNECHET, F. J. & UZE, G. 2006. Interferon-lambda-treated dendritic cells specifically induce proliferation of FOXP3-expressing suppressor T cells. *Blood,* 107, 4417-23.

MERIKA, M. & THANOS, D. 2001. Enhanceosomes. *Curr Opin Genet Dev,* 11, 205-8.

MEYLAN, E., CURRAN, J., HOFMANN, K., MORADPOUR, D., BINDER, M., BARTENSCHLAGER, R. & TSCHOPP, J. 2005. Cardif is an adaptor protein in the RIG-I antiviral pathway and is targeted by hepatitis C virus. *Nature,* 437, 1167-72.

MIBAYASHI, M., MARTINEZ-SOBRIDO, L., LOO, Y. M., CARDENAS, W. B., GALE, M., JR. & GARCIA-SASTRE, A. 2007. Inhibition of retinoic acid-inducible gene I-mediated induction of beta interferon by the NS1 protein of influenza A virus. *J Virol,* 81, 514-24.

MIDGLEY, A. C. & BOWEN, T. 2015. Analysis of human hyaluronan synthase gene transcriptional regulation and downstream hyaluronan cell surface receptor mobility in myofibroblast differentiation. Methods Mol Biol, 1229, 605-18.

MIGNOT, C. C., PIROTTIN, D., FARNIR, F., DE MOFFARTS, B., MOLITOR, C., LEKEUX, P. & ART, T. 2012. Effect of strenuous exercise and ex vivo TLR3 and TLR4 stimulation on inflammatory gene expression in equine pulmonary leukocytes. *Vet Immunol Immunopathol,* 147, 127-35.

MIN, J. Y. & KRUG, R. M. 2006. The primary function of RNA binding by the influenza A virus NS1 protein in infected cells: Inhibiting the 2'-5' oligo (A) synthetase/RNase L pathway. Proc Natl Acad Sci U S A, 103, 7100-5.

MIN, J. Y., LI, S., SEN, G. C. & KRUG, R. M. 2007. A site on the influenza A virus NS1 protein mediates both inhibition of PKR activation and temporal regulation of viral RNA synthesis. *Virology,* 363, 236-43.

MIURA, K. 2008. ERK2-binding domain is required for phosphorylation of EBITEIN1, a potential downstream interactor of ERK2. Biochem Biophys Res Commun, 375, 367-71.

MOAL, V., TEXTORIS, J., BEN AMARA, A., MEHRAJ, V., BERLAND, Y., COLSON, P. & MEGE, J. L. 2013. Chronic hepatitis E virus infection is specifically associated with an interferon-related transcriptional program. J Infect Dis, 207, 125-32.

MOLL, H. P., MAIER, T., ZOMMER, A., LAVOIE, T. & BROSTJAN, C. 2011. The differential activity of interferon-alpha subtypes is consistent among distinct target genes and cell types. Cytokine, 53, 52-9.

MONRO, R. E., CERNA, J. & MARCKER, K. A. 1968. Ribosome-catalyzed peptidyl transfer: substrate specificity at the P-site. Proc Natl Acad Sci U S A, 61, 1042-9.

MORLEY, V. J. & TURNER, P. E. 2017. Dynamics of molecular evolution in RNA virus populations depend on sudden versus gradual environmental change. Evolution, 71, 872-883.

MOUNEIMNE, G. & BRUGGE, J. S. 2007. Tensins: a new switch in cell migration. Dev Cell, 13, 317-9.

MUMFORD, J. A. 1999. The equine influenza surveillance program. *Adv Vet Med,* 41, 379-87.

MUNIR, M. & BERG, M. 2013. The multiple faces of proteinkinase R in antiviral defense. *Virulence,* 4, 85-9.

MUNSHI, N., AGALIOTI, T., LOMVARDAS, S., MERIKA, M., CHEN, G. & THANOS, D. 2001. Coordination of a transcriptional switch by HMGI(Y) acetylation. *Science,* 293, 1133-6.

MURAMOTO, Y., NODA, T., KAWAKAMI, E., AKKINA, R. & KAWAOKA, Y. 2013. Identification of novel influenza A virus proteins translated from PA mRNA. *J Virol,* 87, 2455-62.

MURCIA, P. R., BAILLIE, G. J., DALY, J., ELTON, D., JERVIS, C., MUMFORD, J. A., NEWTON, R., PARRISH, C. R., HOELZER, K., DOUGAN, G., PARKHILL, J., LENNARD, N., ORMOND, D., MOULE, S., WHITWHAM, A., MCCAULEY, J. W., MCKINLEY, T. J., HOLMES, E. C., GRENFELL, B. T. & WOOD, J. L. 2010. Intra- and interhost evolutionary dynamics of equine influenza virus. J Virol, 84, 6943-54.

MURCIA, P. R., WOOD, J. L. & HOLMES, E. C. 2011. Genome-scale evolution and phylodynamics of equine H3N8 influenza A virus. J Virol, 85, 5312-22.

MUZIO, G., MAGGIORA, M., PAIUZZI, E., ORALDI, M. & CANUTO, R. A. 2012. Aldehyde dehydrogenases and cell proliferation. Free Radic Biol Med, 52, 735-46.

MYERS C, W. W. 2006. Equine Influenza Virus. *Clinical Techniques in Equine Practice,* 5, 187-196.

MYONG, S., CUI, S., CORNISH, P. V., KIRCHHOFER, A., GACK, M. U., JUNG, J. U., HOPFNER, K. P. & HA, T. 2009. Cytosolic viral sensor RIG-I is a 5'-triphosphate-dependent translocase on double-stranded RNA. *Science,* 323, 1070-4.

NA, W., LYOO, K. S., YOON, S. W., YEOM, M., KANG, B., MOON, H., KIM, H. K., JEONG, D. G., KIM, J. K. & SONG, D. 2016. Attenuation of the virulence of a recombinant influenza virus expressing the naturally truncated NS gene from an H3N8 equine influenza virus in mice. Vet Res, 47, 115.

NAKHAEI, P., GENIN, P., CIVAS, A. & HISCOTT, J. 2009. RIG-I-like receptors: sensing and responding to RNA virus infection. *Semin Immunol,* 21, 215-22.

NAM, G. H., AHN, K., BAE, J. H., HAN, K., LEE, C. E., PARK, K. D., LEE, S. H., CHO, B. W. & KIM, H. S. 2011. Genomic structure and expression analyses of the PYGM gene in the thoroughbred horse. Zoolog Sci, 28, 276-80.

NEMEROFF, M. E., BARABINO, S. M., LI, Y., KELLER, W. & KRUG, R. M. 1998. Influenza virus NS1 protein interacts with the cellular 30 kDa subunit of CPSF and inhibits 3'end formation of cellular pre-mRNAs. *Mol Cell,* 1, 991-1000.

NEWBY, C. M., SABIN, L. & PEKOSZ, A. 2007. The RNA binding domain of influenza A virus NS1 protein affects secretion of tumor necrosis factor alpha, interleukin-6, and interferon in primary murine tracheal epithelial cells. *J Virol,* 81, 9469-80.

NEWTON, J. R., DALY, J. M., SPENCER, L. & MUMFORD, J. A. 2006. Description of the outbreak of equine influenza (H3N8) in the United Kingdom in 2003, during which recently vaccinated horses in Newmarket developed respiratory disease. *Vet Rec,* 158, 185-92.

NEWTON, J. R., TOWNSEND, H. G., WOOD, J. L., SINCLAIR, R., HANNANT, D. & MUMFORD, J. A. 2000. Immunity to equine influenza: relationship of vaccine-induced antibody in young

Thoroughbred racehorses to protection against field infection with influenza A/equine-2 viruses (H3N8). *Equine Vet J,* 32**,** 65-74.

NGUNJIRI, J. M., LEE, C. W., ALI, A. & MARCUS, P. I. 2012. Influenza virus interferon-inducing particle efficiency is reversed in avian and mammalian cells, and enhanced in cells co-infected with defective-interfering particles. *J Interferon Cytokine Res,* 32**,** 280-5.

NISHIHARA, M., YAMADA, M., NOZAKI, M., NAKAHIRA, K. & YANAGIHARA, I. 2010. Transcriptional regulation of the human establishment of cohesion 1 homolog 2 gene. Biochem Biophys Res Commun, 393, 111-7.

NIU, P., SHABIR, N., KHATUN, A., SEO, B. J., GU, S., LEE, S. M., LIM, S. K., KIM, K. S. & KIM, W. I. 2016. Effect of polymorphisms in the GBP1, Mx1 and CD163 genes on host responses to PRRSV infection in pigs. Vet Microbiol, 182, 187-95.

NOAH, D. L., TWU, K. Y. & KRUG, R. M. 2003. Cellular antiviral responses against influenza A virus are countered at the posttranscriptional level by the viral NS1A protein via its binding to a cellular protein required for the 3' end processing of cellular pre-mRNAS. Virology, 307, 386-95.
NOGALES, A., BAKER, S. F., ORTIZ-RIANO, E., DEWHURST, S., TOPHAM, D. J. & MARTINEZ-SOBRIDO, L. 2014. Influenza A virus attenuation by codon deoptimization of the NS gene for vaccine development. J Virol, 88, 10525-40.

NOLIN, J. D., TULLY, J. E., HOFFMAN, S. M., GUALA, A. S., VAN DER VELDEN, J. L., POYNTER, M. E., VAN DER VLIET, A., ANATHY, V. & JANSSEN-HEININGER, Y. M. 2014. The glutaredoxin/S-glutathionylation axis regulates interleukin-17A-induced proinflammatory responses in lung epithelial cells in association with S-glutathionylation of nuclear factor kappaB family proteins. Free Radic Biol Med, 73, 143-53.

NOGALES, A., CHAUCHE, C., DeDIEGO, M.L., TOPHAM, D.J., PARRISH C.R., MURCIA P.R., MARTINEZ-SOBRIDO, L. 2017. The K186E Amino Acid Substitution in the Canine Influenza Virus H3N8 NS1 Protein Restores Its Ability To Inhibit Host Gene Expression. J Virol, 91(22).

NOYAMA, Y., OKANO, M., FUJIWARA, T., KARIYA, S., HIGAKI, T., HARUNA, T., MAKIHARA, S. I., KANAI, K., KOYAMA, T., TANIGUCHI, M., ISHITOYA, J. I., KANDA, A., KOBAYASHI, Y., ASAKO, M., TOMODA, K. & NISHIZAKI, K. 2017. IL-22/IL-22R1 signaling regulates the pathophysiology of chronic rhinosinusitis with nasal polyps via alteration of MUC1 expression. Allergol Int, 66, 42-51.

O'CONNELL, B. C., ADAMSON, B., LYDEARD, J. R., SOWA, M. E., CICCIA, A., BREDEMEYER, A. L., SCHLABACH, M., GYGI, S. P., ELLEDGE, S. J. & HARPER, J. W. 2010. A genome-wide camptothecin sensitivity screen identifies a mammalian MMS22L-NFKBIL2 complex required for genomic stability. Mol Cell, 40, 645-57.

OBENAUER, J. C., DENSON, J., MEHTA, P. K., SU, X., MUKATIRA, S., FINKELSTEIN, D. B., XU, X., WANG, J., MA, J., FAN, Y., RAKESTRAW, K. M., WEBSTER, R. G., HOFFMANN, E., KRAUSS, S., ZHENG, J., ZHANG, Z. & NAEVE, C. W. 2006. Large-scale sequence analysis of avian influenza isolates. Science, 311, 1576-80.

OKADA, M., CHEESEMAN, I. M., HORI, T., OKAWA, K., MCLEOD, I. X., YATES, J. R., 3RD, DESAI, A. & FUKAGAWA, T. 2006. The CENP-H-I complex is required for the efficient incorporation of newly synthesized CENP-A into centromeres. Nat Cell Biol, 8, 446-57.

OKUMOTO, K., KAMETANI, Y. & FUJIKI, Y. 2011. Two proteases, trypsin domain-containing 1 (Tysnd1) and peroxisomal lon protease (PsLon), cooperatively regulate fatty acid beta-oxidation in peroxisomal matrix. J Biol Chem, 286, 44367-79.

OLGUIN PERGLIONE, C., GOLEMBA, M. D., TORRES, C. & BARRANDEGUY, M. 2016. Molecular Epidemiology and Spatio-Temporal Dynamics of the H3N8 Equine Influenza Virus in South America. *Pathogens,* 5.

ONOGUCHI, K., YONEYAMA, M., TAKEMURA, A., AKIRA, S., TANIGUCHI, T., NAMIKI, H. & FUJITA, T. 2007. Viral infections activate types I and III interferon genes through a common mechanism. *J Biol Chem,* 282**,** 7576-81.

ONOMOTO, K., JOGI, M., YOO, J. S., NARITA, R., MORIMOTO, S., TAKEMURA, A., SAMBHARA, S., KAWAGUCHI, A., OSARI, S., NAGATA, K., MATSUMIYA, T., NAMIKI, H., YONEYAMA, M. & FUJITA, T. 2012. Critical role of an antiviral stress granule containing RIG-I and PKR in viral detection and innate immunity. *PLoS One,* 7**,** e43031.

OOI, A., WONG, A., ESAU, L., LEMTIRI-CHLIEH, F. & GEHRING, C. 2016. A Guide to Transient Expression of Membrane Proteins in HEK-293 Cells for Functional Characterization. Front Physiol, 7, 300.

OPITZ, B., REJAIBI, A., DAUBER, B., ECKHARD, J., VINZING, M., SCHMECK, B., HIPPENSTIEL, S., SUTTORP, N. & WOLFF, T. 2007. IFNbeta induction by influenza A virus is mediated by RIG-I which is regulated by the viral NS1 protein. *Cell Microbiol,* 9**,** 930-8.

OSHIUMI, H., MIYASHITA, M., MATSUMOTO, M. & SEYA, T. 2013. A distinct role of Riplet-mediated K63-Linked polyubiquitination of the RIG-I repressor domain in human antiviral innate immune responses. *PLoS Pathog,* 9**,** e1003533.

OSTERLUND, P. I., PIETILA, T. E., VECKMAN, V., KOTENKO, S. V. & JULKUNEN, I. 2007. IFN regulatory factor family members differentially regulate the expression of type III IFN (IFN-lambda) genes. *J Immunol,* 179**,** 3434-42.

OTERA, H., ISHIDA, T., NISHIUMA, T., KOBAYASHI, K., KOTANI, Y., YASUDA, T., KUNDU, R. K., QUERTERMOUS, T., HIRATA, K. & NISHIMURA, Y. 2009. Targeted inactivation of endothelial lipase attenuates lung allergic inflammation through raising plasma HDL level and inhibiting eosinophil infiltration. Am J Physiol Lung Cell Mol Physiol, 296, L594-602.

PACCAUD, M. F. & PACCAUD, M. 1967. Influenza A-equi-2 viruses isolated in Switzerland, 1965-1966. A comparative antigenic study. *Pathol Microbiol (Basel),* 30**,** 999-1006.

PAL, S., SANTOS, A., ROSAS, J. M., ORTIZ-GUZMAN, J. & ROSAS-ACOSTA, G. 2011. Influenza A virus interacts extensively with the cellular SUMOylation system during infection. *Virus Res,* 158**,** 12-27.

PARK, A. W., WOOD, J. L., DALY, J. M., NEWTON, J. R., GLASS, K., HENLEY, W., MUMFORD, J. A. & GRENFELL, B. T. 2004. The effects of strain heterology on the epidemiology of equine influenza in a vaccinated population. *Proc Biol Sci,* 271**,** 1547-55.

PARK, S. J., KUMAR, M., KWON, H. I., SEONG, R. K., HAN, K., SONG, J. M., KIM, C. J., CHOI, Y. K. & SHIN, O. S. 2015. Dynamic changes in host gene expression associated with H5N8 avian influenza virus infection in mice. Sci Rep, 5, 16512.

PARRISH, C. R. & KAWAOKA, Y. 2005. The origins of new pandemic viruses: the acquisition of new host ranges by canine parvovirus and influenza A viruses. *Annu Rev Microbiol,* 59**,** 553-86.

PARRISH, C. R., MURCIA, P. R. & HOLMES, E. C. 2015. Influenza virus reservoirs and intermediate hosts: dogs, horses, and new possibilities for influenza virus exposure of humans. *Journal of virology,* 89**,** 2990-4.

PAZ, S., SUN, Q., NAKHAEI, P., ROMIEU-MOUREZ, R., GOUBAU, D., JULKUNEN, I., LIN, R. & HISCOTT, J. 2006. Induction of IRF-3 and IRF-7 phosphorylation following activation of the RIG-I pathway. *Cell Mol Biol (Noisy-le-grand),* 52**,** 17-28.

PECHHOLD, K., PATTERSON, N. B., CRAIGHEAD, N., LEE, K. P., JUNE, C. H. & HARLAN, D. M. 1997. Inflammatory cytokines IFN-gamma plus TNF-alpha induce regulated expression of CD80 (B7-1) but not CD86 (B7-2) on murine fibroblasts. J Immunol, 158, 4921-9.

PEDERSEN, B. K. & TOFT, A. D. 2000. Effects of exercise on lymphocytes and cytokines. *Br J Sports Med,* 34, 246-51.

PEI, J. J., AN, W. L., ZHOU, X. W., NISHIMURA, T., NORBERG, J., BENEDIKZ, E., GOTZ, J. & WINBLAD, B. 2006. P70 S6 kinase mediates tau phosphorylation and synthesis. FEBS Lett, 580, 107-14.

PEREZ-CIDONCHA, M., KILLIP, M. J., OLIVEROS, J. C., ASENSIO, V. J., FERNANDEZ, Y., BENGOECHEA, J. A., RANDALL, R. E. & ORTIN, J. 2014. An unbiased genetic screen reveals the polygenic nature of the influenza virus anti-interferon response. *J Virol,* 88, 4632-46.

PERGLIONE, C. O., GILDEA, S., RIMONDI, A., MINO, S., VISSANI, A., CAROSSINO, M., CULLINANE, A. & BARRANDEGUY, M. 2016. Epidemiological and virological findings during multiple outbreaks of equine influenza in South America in 2012. *Influenza Other Respir Viruses,* 10, 37-46.

PICHLMAIR, A., SCHULZ, O., TAN, C. P., NASLUND, T. I., LILJESTROM, P., WEBER, F. & REIS E SOUSA, C. 2006. RIG-I-mediated antiviral responses to single-stranded RNA bearing 5'-phosphates. Science, 314, 997-1001.

PINTO, A. K., WILLIAMS, G. D., SZRETTER, K. J., WHITE, J. P., PROENCA-MODENA, J. L., LIU, G., OLEJNIK, J., BRIEN, J. D., EBIHARA, H., MUHLBERGER, E., AMARASINGHE, G., DIAMOND, M. S. & BOON, A. C. 2015. Human and Murine IFIT1 Proteins Do Not Restrict Infection of Negative-Sense RNA Viruses of the Orthomyxoviridae, Bunyaviridae, and Filoviridae Families. J Virol, 89, 9465-76.

PIRILLO, A., CATAPANO, A. L. & NORATA, G. D. 2015. HDL in infectious diseases and sepsis. Handb Exp Pharmacol, 224, 483-508.

PIWKO, W., OLMA, M. H., HELD, M., BIANCO, J. N., PEDRIOLI, P. G., HOFMANN, K., PASERO, P., GERLICH, D. W. & PETER, M. 2010. RNAi-based screening identifies the Mms22L-Nfkbil2 complex as a novel regulator of DNA replication in human cells. EMBO J, 29, 4210-22.

POCIASK, D. A., ROBINSON, K. M., CHEN, K., MCHUGH, K. J., CLAY, M. E., HUANG, G. T., BENOS, P. V., JANSSEN-HEININGER, Y. M. W., KOLLS, J. K., ANATHY, V. & ALCORN, J. F. 2017. Epigenetic and Transcriptomic Regulation of Lung Repair during Recovery from Influenza Infection. Am J Pathol, 187, 851-863.

POON, L. L., PRITLOVE, D. C., FODOR, E. & BROWNLEE, G. G. 1999. Direct evidence that the poly(A) tail of influenza A virus mRNA is synthesized by reiterative copying of a U track in the virion RNA template. *J Virol,* 73, 3473-6.

PRAKASH, A., SMITH, E., LEE, C. K. & LEVY, D. E. 2005. Tissue-specific positive feedback requirements for production of type I interferon following virus infection. *J Biol Chem,* 280, 18651-7.

PRECIOUS, B., CHILDS, K., FITZPATRICK-SWALLOW, V., GOODBOURN, S. & RANDALL, R. E. 2005. Simian virus 5 V protein acts as an adaptor, linking DDB1 to STAT2, to facilitate the ubiquitination of STAT1. *J Virol,* 79, 13434-41.

PRIVALSKY, M. L. & PENHOET, E. E. 1978. Influenza virus proteins: identity, synthesis, and modification analyzed by two-dimensional gel electrophoresis. *Proc Natl Acad Sci U S A,* 75, 3625-9.

PU, S. Y., YU, Q., WU, H., JIANG, J. J., CHEN, X. Q., HE, Y. H. & KONG, Q. P. 2017. ERCC6L, a DNA helicase, is involved in cell proliferation and associated with survival and progress in breast and kidney cancers. Oncotarget, 8, 42116-42124.

PUGH, D. J., AB, E., FARO, A., LUTYA, P. T., HOFFMANN, E. & REES, D. J. 2006. DWNN, a novel ubiquitin-like domain, implicates RBBP6 in mRNA processing and ubiquitin-like pathways. BMC Struct Biol, 6, 1.

QASHQARI, H., AL-MARS, A., CHAUDHARY, A., ABUZENADAH, A., DAMANHOURI, G., ALQAHTANI, M., MAHMOUD, M., EL SAYED ZAKI, M., FATIMA, K. & QADRI, I. 2013. Understanding the molecular mechanism(s) of hepatitis C virus (HCV) induced interferon resistance. Infect Genet Evol, 19, 113-9.

QI, T., GUO, W., HUANG, W. Q., LI, H. M., ZHAO, L. P., DAI, L. L., HE, N., HAO, X. F. & XIANG, W. H. 2010. Genetic evolution of equine influenza viruses isolated in China. Arch Virol, 155, 1425-32.

QIAN, X. Y., ALONSO-CAPLEN, F. & KRUG, R. M. 1994. Two functional domains of the influenza virus NS1 protein are required for regulation of nuclear export of mRNA. *J Virol,* 68**,** 2433-41.
QIU, Y. & KRUG, R. M. 1994. The influenza virus NS1 protein is a poly(A)-binding protein that inhibits nuclear export of mRNAs containing poly(A). *J Virol,* 68**,** 2425-32.

QUANTIUS, J., SCHMOLDT, C., VAZQUEZ-ARMENDARIZ, A. I., BECKER, C., EL AGHA, E., WILHELM, J., MORTY, R. E., VADASZ, I., MAYER, K., GATTENLOEHNER, S., FINK, L., MATROSOVICH, M., LI, X., SEEGER, W., LOHMEYER, J., BELLUSCI, S. & HEROLD, S. 2016. Influenza Virus Infects Epithelial Stem/Progenitor Cells of the Distal Lung: Impact on Fgfr2b-Driven Epithelial Repair. PLoS Pathog, 12, e1005544.
QUINLIVAN, M., ZAMARIN, D., GARCIA-SASTRE, A., CULLINANE, A., CHAMBERS, T. & PALESE, P. 2005. Attenuation of equine influenza viruses through truncations of the NS1 protein. *J Virol,* 79**,** 8431-9.

RAJSBAUM, R., ALBRECHT, R. A., WANG, M. K., MAHARAJ, N. P., VERSTEEG, G. A., NISTAL-VILLAN, E., GARCIA-SASTRE, A. & GACK, M. U. 2012. Species-specific inhibition of RIG-I ubiquitination and IFN induction by the influenza A virus NS1 protein. PLoS Pathog, 8, e1003059.
RANDALL, R. E. & GOODBOURN, S. 2008. Interferons and viruses: an interplay between induction, signalling, antiviral responses and virus countermeasures. *The Journal of general virology,* 89**,** 1-47.

RASCHLE, M., SMEENK, G., HANSEN, R. K., TEMU, T., OKA, Y., HEIN, M. Y., NAGARAJ, N., LONG, D. T., WALTER, J. C., HOFMANN, K., STORCHOVA, Z., COX, J., BEKKER-JENSEN, S., MAILAND, N. & MANN, M. 2015. DNA repair. Proteomics reveals dynamic assembly of repair complexes during bypass of DNA cross-links. Science, 348, 1253671.

RASS, U., COMPTON, S. A., MATOS, J., SINGLETON, M. R., IP, S. C., BLANCO, M. G., GRIFFITH, J. D. & WEST, S. C. 2010. Mechanism of Holliday junction resolution by the human GEN1 protein. Genes Dev, 24, 1559-69.

RATINIER, M., SHAW, A. E., BARRY, G., GU, Q., DI GIALLEONARDO, L., JANOWICZ, A., VARELA, M., RANDALL, R. E., CAPORALE, M. & PALMARINI, M. 2016. Bluetongue Virus NS4 Protein Is an Interferon Antagonist and a Determinant of Virus Virulence. Journal of virology, 90, 5427-39.

REHWINKEL, J., TAN, C. P., GOUBAU, D., SCHULZ, O., PICHLMAIR, A., BIER, K., ROBB, N., VREEDE, F., BARCLAY, W., FODOR, E. & REIS E SOUSA, C. 2010. RIG-I detects viral genomic RNA during negative-strand RNA virus infection. *Cell,* 140**,** 397-408.

REICH, S., GUILLIGAY, D., PFLUG, A., MALET, H., BERGER, I., CREPIN, T., HART, D., LUNARDI, T., NANAO, M., RUIGROK, R. W. & CUSACK, S. 2014. Structural insight into cap-snatching and RNA synthesis by influenza polymerase. *Nature,* 516**,** 361-6.

RESA-INFANTE, P., JORBA, N., COLOMA, R. & ORTIN, J. 2011. The influenza virus RNA synthesis machine: advances in its structure and function. *RNA Biol,* 8**,** 207-15.

REYNAERT, N. L., WOUTERS, E. F. & JANSSEN-HEININGER, Y. M. 2007. Modulation of glutaredoxin-1 expression in a mouse model of allergic airway disease. Am J Respir Cell Mol Biol, 36, 147-51.

RIVAILLER, P., PERRY, I. A., JANG, Y., DAVIS, C. T., CHEN, L. M., DUBOVI, E. J. & DONIS, R. O. 2010. Evolution of canine and equine influenza (H3N8) viruses co-circulating between 2005 and 2008. *Virology,* 408**,** 71-9.

RIVERA, J., PROIA, R. L. & OLIVERA, A. 2008. The alliance of sphingosine-1-phosphate and its receptors in immunity. Nat Rev Immunol, 8, 753-63.

ROBB, N. C., CHASE, G., BIER, K., VREEDE, F. T., SHAW, P. C., NAFFAKH, N., SCHWEMMLE, M. & FODOR, E. 2011. The influenza A virus NS1 protein interacts with the nucleoprotein of viral ribonucleoprotein complexes. *J Virol,* 85**,** 5228-31.

ROSSMAN, J. S. & LAMB, R. A. 2011. Influenza virus assembly and budding. *Virology,* 411**,** 229-36.
ROTHENFUSSER, S., GOUTAGNY, N., DIPERNA, G., GONG, M., MONKS, B. G., SCHOENEMEYER, A., YAMAMOTO, M., AKIRA, S. & FITZGERALD, K. A. 2005. The RNA helicase Lgp2 inhibits TLR-independent sensing of viral replication by retinoic acid-inducible gene-I. *J Immunol,* 175**,** 5260-8.

ROTHWELL, C., LEBRETON, A., YOUNG NG, C., LIM, J. Y., LIU, W., VASUDEVAN, S., LABOW, M., GU, F. & GAITHER, L. A. 2009. Cholesterol biosynthesis modulation regulates dengue viral replication. Virology, 389, 8-19.
SAIRA, K., LIN, X., DEPASSE, J. V., HALPIN, R., TWADDLE, A., STOCKWELL, T., ANGUS, B., COZZI-LEPRI, A., DELFINO, M., DUGAN, V., DWYER, D. E., FREIBERG, M., HORBAN, A., LOSSO, M., LYNFIELD, R., WENTWORTH, D. N., HOLMES, E. C., DAVEY, R., WENTWORTH, D. E., GHEDIN, E., GROUP, I. F. S. & GROUP, I. F. S. 2013. Sequence analysis of in vivo defective interfering-like RNA of influenza A H1N1 pandemic virus. *J Virol,* 87**,** 8064-74.

SAKABE, S., OZAWA, M., TAKANO, R., IWASTUKI-HORIMOTO, K. & KAWAOKA, Y. 2011. Mutations in PA, NP, and HA of a pandemic (H1N1) 2009 influenza virus contribute to its adaptation to mice. Virus research, 158, 124-9.

SANTOS, A., PAL, S., CHACON, J., MERAZ, K., GONZALEZ, J., PRIETO, K. & ROSAS-ACOSTA, G. 2013. SUMOylation affects the interferon blocking activity of the influenza A nonstructural protein NS1 without affecting its stability or cellular localization. *J Virol,* 87**,** 5602-20.

SARKAR, S. N., PETERS, K. L., ELCO, C. P., SAKAMOTO, S., PAL, S. & SEN, G. C. 2004. Novel roles of TLR3 tyrosine phosphorylation and PI3 kinase in double-stranded RNA signaling. Nat Struct Mol Biol, 11, 1060-7.

SATO, H., OSHIUMI, H., TAKAKI, H., HIKONO, H. & SEYA, T. 2015. Evolution of the DEAD box helicase family in chicken: chickens have no DHX9 ortholog. Microbiol Immunol, 59, 633-40.
SATTERLY, N., TSAI, P. L., VAN DEURSEN, J., NUSSENZVEIG, D. R., WANG, Y., FARIA, P. A., LEVAY, A., LEVY, D. E. & FONTOURA, B. M. 2007. Influenza virus targets the mRNA export machinery and the nuclear pore complex. *Proc Natl Acad Sci U S A,* 104**,** 1853-8.

SCHAALE, K., PETERS, K. M., MURTHY, A. M., FRITZSCHE, A. K., PHAN, M. D., TOTSIKA, M., ROBERTSON, A. A., NICHOLS, K. B., COOPER, M. A., STACEY, K. J., ULETT, G. C., SCHRODER, K., SCHEMBRI, M. A. & SWEET, M. J. 2016. Strain- and host species-specific inflammasome activation, IL-1beta release, and cell death in macrophages infected with uropathogenic Escherichia coli. *Mucosal Immunol,* 9**,** 124-36.

SCHNEIDER, W. M., CHEVILLOTTE, M. D. & RICE, C. M. 2014. Interferon-stimulated genes: a complex web of host defenses. *Annu Rev Immunol,* 32**,** 513-45.

SCHOGGINS, J. W. & RANDALL, G. 2013. Lipids in innate antiviral defense. *Cell Host Microbe,* 14**,** 379-85.

SCHOLTENS, R. G., STEELE, J. H., DOWDLE, W. R., YARBROUGH, W. B. & ROBINSON, R. Q. 1964. U.S. Epizootic of Equine Influenza, 1963. *Public Health Rep,* 79**,** 393-402.

SCHOLTISSEK, C. & VON HOYNINGEN-HUENE, V. 1980. Genetic relatedness of the gene which codes for the nonstructural (NS) protein of different influenza A strains. *Virology,* 102**,** 13-20.
SCHOPF, F. H., BIEBL, M. M. & BUCHNER, J. 2017. The HSP90 chaperone machinery. Nat Rev Mol Cell Biol, 18, 345-360.

SCHRODER, E., GEBEL, L., EREMEEV, A. A., MORGNER, J., GRUM, D., KNAUER, S. K., BAYER, P. & MUELLER, J. W. 2012. Human PAPS synthase isoforms are dynamically regulated enzymes with access to nucleus and cytoplasm. PLoS One, 7, e29559.

SCHULTZ-CHERRY, S., DYBDAHL-SISSOKO, N., NEUMANN, G., KAWAOKA, Y. & HINSHAW, V. S. 2001. Influenza virus ns1 protein induces apoptosis in cultured cells. J Virol, 75, 7875-81.

SEITZ, C., FRENSING, T., HOPER, D., KOCHS, G. & REICHL, U. 2010. High yields of influenza A virus in Madin-Darby canine kidney cells are promoted by an insufficient interferon-induced antiviral state. J Gen Virol, 91, 1754-63.

SELIGE, J., HATZELMANN, A. & DUNKERN, T. 2011. The differential impact of PDE4 subtypes in human lung fibroblasts on cytokine-induced proliferation and myofibroblast conversion. J Cell Physiol, 226, 1970-80.

SEO, S. H., HOFFMANN, E. & WEBSTER, R. G. 2002. Lethal H5N1 influenza viruses escape host anti-viral cytokine responses. *Nat Med,* 8**,** 950-4.

SERQUINA, A. K. P., KAMBACH, D. M., SARKER, O. & ZIEGELBAUER, J. M. 2017. Viral MicroRNAs Repress the Cholesterol Pathway, and 25-Hydroxycholesterol Inhibits Infection. MBio, 8.

SETH, R. B., SUN, L., EA, C. K. & CHEN, Z. J. 2005. Identification and characterization of MAVS, a mitochondrial antiviral signaling protein that activates NF-kappaB and IRF 3. *Cell,* 122**,** 669-82.

SGORBISSA, A. & BRANCOLINI, C. 2012. IFNs, ISGylation and cancer: Cui prodest? *Cytokine Growth Factor Rev,* 23**,** 307-14.

SHAPIRO, G. I., GURNEY, T., JR. & KRUG, R. M. 1987. Influenza virus gene expression: control mechanisms at early and late times of infection and nuclear-cytoplasmic transport of virus-specific RNAs. *J Virol,* 61**,** 764-73.

SHARMA, K., TRIPATHI, S., RANJAN, P., KUMAR, P., GARTEN, R., DEYDE, V., KATZ, J. M., COX, N. J., LAL, R. B., SAMBHARA, S. & LAL, S. K. 2011. Influenza A virus nucleoprotein exploits Hsp40 to inhibit PKR activation. *PLoS One,* 6**,** e20215.

SHELTON, H., SMITH, M., HARTGROVES, L., STILWELL, P., ROBERTS, K., JOHNSON, B. & BARCLAY, W. 2012. An influenza reassortant with polymerase of pH1N1 and NS gene of H3N2 influenza A virus is attenuated in vivo. J Gen Virol, 93, 998-1006.

SHEN, X., LI, J., HU, P. P., WADDELL, D., ZHANG, J. & WANG, X. F. 2001. The activity of guanine exchange factor NET1 is essential for transforming growth factor-beta-mediated stress fiber formation. J Biol Chem, 276, 15362-8.

SHI, L., DING, W., LI, D., WANG, Z., JIANG, H., ZHANG, J. & TANG, C. 2006. Proliferation and anti-apoptotic effects of human urotensin II on human endothelial cells. Atherosclerosis, 188, 260-4.

SHIMIZU, K., IGUCHI, A., GOMYOU, R. & ONO, Y. 1999. Influenza virus inhibits cleavage of the HSP70 pre-mRNAs at the polyadenylation site. Virology, 254, 213-9.

SHIN, Y. K., LI, Y., LIU, Q., ANDERSON, D. H., BABIUK, L. A. & ZHOU, Y. 2007a. SH3 binding motif 1 in influenza A virus NS1 protein is essential for PI3K/Akt signaling pathway activation. J Virol, 81, 12730-9.

SHIN, Y. K., LIU, Q., TIKOO, S. K., BABIUK, L. A. & ZHOU, Y. 2007b. Effect of the phosphatidylinositol 3-kinase/Akt pathway on influenza A virus propagation. J Gen Virol, 88, 942-50.

SHIN, Y. K., LIU, Q., TIKOO, S. K., BABIUK, L. A. & ZHOU, Y. 2007c. Influenza A virus NS1 protein activates the phosphatidylinositol 3-kinase (PI3K)/Akt pathway by direct interaction with the p85 subunit of PI3K. J Gen Virol, 88, 13-8.

SHIOMI, Y. & NISHITANI, H. 2013. Alternative replication factor C protein, Elg1, maintains chromosome stability by regulating PCNA levels on chromatin. Genes Cells, 18, 946-59.

SIKDAR, N., BANERJEE, S., LEE, K. Y., WINCOVITCH, S., PAK, E., NAKANISHI, K., JASIN, M., DUTRA, A. & MYUNG, K. 2009. DNA damage responses by human ELG1 in S phase are important to maintain genomic integrity. Cell Cycle, 8, 3199-207.

SILVERMAN, R. H. 2007. Viral encounters with 2',5'-oligoadenylate synthetase and RNase L during the interferon antiviral response. *J Virol,* 81, 12720-9.

SKEHEL, J. J. 1973. Early polypeptide synthesis in influenza virus-infected cells. *Virology,* 56, 394-9.

SMYTH, G. B., DAGLEY, K. & TAINSH, J. 2011. Insights into the economic consequences of the 2007 equine influenza outbreak in Australia. *Aust Vet J,* 89 Suppl 1, 151-8.

SOBOLL HUSSEY, G., HUSSEY, S. B., WAGNER, B., HOROHOV, D. W., VAN DE WALLE, G. R., OSTERRIEDER, N., GOEHRING, L. S., RAO, S. & LUNN, D. P. 2011. Evaluation of immune responses following infection of ponies with an EHV-1 ORF1/2 deletion mutant. Vet Res, 42, 23.

SOLORZANO, A., WEBBY, R. J., LAGER, K. M., JANKE, B. H., GARCIA-SASTRE, A. & RICHT, J. A. 2005. Mutations in the NS1 protein of swine influenza virus impair anti-interferon activity and confer attenuation in pigs. J Virol, 79, 7535-43.

SOUBIES, S. M., VOLMER, C., CROVILLE, G., LOUPIAS, J., PERALTA, B., COSTES, P., LACROUX, C., GUERIN, J. L. & VOLMER, R. 2010. Species-specific contribution of the four C-terminal amino acids of influenza A virus NS1 protein to virulence. *J Virol,* 84, 6733-47.

SOVINOVA, O. & LUDVIK, J. 1959. Electron microscopic study of the influenza virus A-equi-Praha/56. *Acta Virol,* 3, 59-60.

STACK, J. C., MURCIA, P. R., GRENFELL, B. T., WOOD, J. L. & HOLMES, E. C. 2013. Inferring the inter-host transmission of influenza A virus using patterns of intra-host genetic variation. *Proc Biol Sci,* 280, 20122173.

STARCK, S. R., GREEN, H. M., ALBEROLA-ILA, J. & ROBERTS, R. W. 2004. A general approach to detect protein expression in vivo using fluorescent puromycin conjugates. Chem Biol, 11, 999-1008.

STAROKADOMSKYY, P., GLUCK, N., LI, H., CHEN, B., WALLIS, M., MAINE, G. N., MAO, X., ZAIDI, I. W., HEIN, M. Y., MCDONALD, F. J., LENZNER, S., ZECHA, A., ROPERS, H. H., KUSS, A. W., MCGAUGHRAN, J., GECZ, J. & BURSTEIN, E. 2013. CCDC22 deficiency in humans blunts activation of proinflammatory NF-kappaB signaling. J Clin Invest, 123, 2244-56.

STEIDLE, S., MARTINEZ-SOBRIDO, L., MORDSTEIN, M., LIENENKLAUS, S., GARCIA-SASTRE, A., STAHELI, P. & KOCHS, G. 2010. Glycine 184 in nonstructural protein NS1 determines the virulence of influenza A virus strain PR8 without affecting the host interferon response. J Virol, 84, 12761-70.

STEINHAUER, D. A. & SKEHEL, J. J. 2002. Genetics of influenza viruses. *Annu Rev Genet,* 36, 305-32.

STEWART, C. E., RANDALL, R. E. & ADAMSON, C. S. 2014. Inhibitors of the interferon response enhance virus replication in vitro. PloS one, 9, e112014.

STRAHLE, L., GARCIN, D. & KOLAKOFSKY, D. 2006. Sendai virus defective-interfering genomes and the activation of interferon-beta. *Virology,* 351, 101-11.

STUBBS, T. M. & TE VELTHUIS, A. J. 2014. The RNA-dependent RNA polymerase of the influenza A virus. *Future Virol,* 9, 863-876.

SUAREZ, D. L. & PERDUE, M. L. 1998. Multiple alignment comparison of the non-structural genes of influenza A viruses. Virus Res, 54, 59-69.

SUGIYAMA, K., KAWAGUCHI, A., OKUWAKI, M. & NAGATA, K. 2015. pp32 and APRIL are host cell-derived regulators of influenza virus RNA synthesis from cRNA. *Elife,* 4.

SUNAGA, H., MATSUI, H., UENO, M., MAENO, T., ISO, T., SYAMSUNARNO, M. R., ANJO, S., MATSUZAKA, T., SHIMANO, H., YOKOYAMA, T. & KURABAYASHI, M. 2013. Deranged fatty acid composition causes pulmonary fibrosis in Elovl6-deficient mice. Nat Commun, 4, 2563.

SWIECKI, M., SCHEAFFER, S. M., ALLAIRE, M., FREMONT, D. H., COLONNA, M. & BRETT, T. J. 2011. Structural and biophysical analysis of BST-2/tetherin ectodomains reveals an evolutionary conserved design to inhibit virus release. *J Biol Chem,* 286, 2987-97.

SZE, A., OLAGNIER, D., LIN, R., VAN GREVENYNGHE, J. & HISCOTT, J. 2013. SAMHD1 host restriction factor: a link with innate immune sensing of retrovirus infection. J Mol Biol, 425, 4981-94.

SZEWCZYK, B., BIENKOWSKA-SZEWCZYK, K. & KROL, E. 2014. Introduction to molecular biology of influenza a viruses. *Acta Biochim Pol,* 61, 397-401.

TABUCHI, I. 2003. Next-generation protein-handling method: puromycin analogue technology. Biochem Biophys Res Commun, 305, 1-5.

TAKATA, H., MATSUNAGA, S., MORIMOTO, A., ONO-MANIWA, R., UCHIYAMA, S. & FUKUI, K. 2007. H1.X with different properties from other linker histones is required for mitotic progression. FEBS Lett, 581, 3783-8.

TAKIZAWA, T., FUKUDA, R., MIYAWAKI, T., OHASHI, K. & NAKANISHI, Y. 1995. Activation of the apoptotic Fas antigen-encoding gene upon influenza virus infection involving spontaneously produced beta-interferon. Virology, 209, 288-96.

TALIPOV, M. R., NAYAK, J., LEPLEY, M., BONGARD, R. D., SEM, D. S., RAMCHANDRAN, R. & RATHORE, R. 2016. Critical Role of the Secondary Binding Pocket in Modulating the Enzymatic Activity of DUSP5 toward Phosphorylated ERKs. Biochemistry, 55, 6187-6195.

TALON, J., HORVATH, C. M., POLLEY, R., BASLER, C. F., MUSTER, T., PALESE, P. & GARCIA-SASTRE, A. 2000. Activation of interferon regulatory factor 3 is inhibited by the influenza A virus NS1 protein. J Virol, 74, 7989-96.

TALON, J., SALVATORE, M., O'NEILL, R. E., NAKAYA, Y., ZHENG, H., MUSTER, T., GARCIA-SASTRE, A. & PALESE, P. 2000b. Influenza A and B viruses expressing altered NS1 proteins: A vaccine approach. Proc Natl Acad Sci U S A, 97, 4309-14.

TAN, S. L. & KATZE, M. G. 1998. Biochemical and genetic evidence for complex formation between the influenza A virus NS1 protein and the interferon-induced PKR protein kinase. J Interferon Cytokine Res, 18, 757-66.

TANG, X., GAO, J. S., GUAN, Y. J., MCLANE, K. E., YUAN, Z. L., RAMRATNAM, B. & CHIN, Y. E. 2007. Acetylation-dependent signal transduction for type I interferon receptor. *Cell,* 131, 93-105.
TANG, Y., ZHONG, G., ZHU, L., LIU, X., SHAN, Y., FENG, H., BU, Z., CHEN, H. & WANG, C. 2010. Herc5 attenuates influenza A virus by catalyzing ISGylation of viral NS1 protein. *J Immunol,* 184, 5777-90.

TAPIA, K., KIM, W. K., SUN, Y., MERCADO-LOPEZ, X., DUNAY, E., WISE, M., ADU, M. & LOPEZ, C. B. 2013. Defective viral genomes arising in vivo provide critical danger signals for the triggering of lung antiviral immunity. *PLoS Pathog,* 9, e1003703.

TATEISHI, Y., ARIYOSHI, M., IGARASHI, R., HARA, H., MIZUGUCHI, K., SETO, A., NAKAI, A., KOKUBO, T., TOCHIO, H. & SHIRAKAWA, M. 2009. Molecular basis for SUMOylation-dependent regulation of DNA binding activity of heat shock factor 2. J Biol Chem, 284, 2435-47.

TEMPLE, R., ALLEN, E., FORDHAM, J., PHIPPS, S., SCHNEIDER, H. C., LINDAUER, K., HAYES, I., LOCKEY, J., POLLOCK, K. & JUPP, R. 2001. Microarray analysis of eosinophils reveals a number of candidate survival and apoptosis genes. Am J Respir Cell Mol Biol, 25, 425-33.

THOMAS, G., BROWN, A. L. & BROWN, J. M. 2014. In vivo metabolite profiling as a means to identify uncharacterized lipase function: recent success stories within the alpha beta hydrolase domain (ABHD) enzyme family. Biochim Biophys Acta, 1841, 1097-101.
THOMAS, P. & SMART, T. G. 2005. HEK293 cell line: a vehicle for the expression of recombinant proteins. J Pharmacol Toxicol Methods, 51, 187-200.

TIAN, B. & MANLEY, J. L. 2017. Alternative polyadenylation of mRNA precursors. Nat Rev Mol Cell Biol, 18, 18-30.

TIAN, E., STEVENS, S. R., GUAN, Y., SPRINGER, D. A., ANDERSON, S. A., STAROST, M. F., PATEL, V., TEN HAGEN, K. G. & TABAK, L. A. 2015. Galnt1 is required for normal heart valve development and cardiac function. PLoS One, 10, e0115861.

TIOLLIER, E., GOMEZ-MERINO, D., BURNAT, P., JOUANIN, J. C., BOURRILHON, C., FILAIRE, E., GUEZENNEC, C. Y. & CHENNAOUI, M. 2005. Intense training: mucosal immunity and incidence of respiratory infections. *Eur J Appl Physiol,* 93**,** 421-8.

TOGAYACHI, A., KOZONO, Y., ISHIDA, H., ABE, S., SUZUKI, N., TSUNODA, Y., HAGIWARA, K., KUNO, A., OHKURA, T., SATO, N., SATO, T., HIRABAYASHI, J., IKEHARA, Y., TACHIBANA, K. & NARIMATSU, H. 2007. Polylactosamine on glycoproteins influences basal levels of lymphocyte and macrophage activation. Proc Natl Acad Sci U S A, 104, 15829-34.

TRAPNELL, C., HENDRICKSON, D. G., SAUVAGEAU, M., GOFF, L., RINN, J. L. & PACHTER, L. 2013. Differential analysis of gene regulation at transcript resolution with RNA-seq. Nature biotechnology, 31, 46-53.

TURKINGTON, H. L., JUOZAPAITIS, M., KERRY, P. S., AYDILLO, T., AYLLON, J., GARCIA-SASTRE, A., SCHWEMMLE, M. & HALE, B. G. 2015. Novel Bat Influenza Virus NS1 Proteins Bind Double-Stranded RNA and Antagonize Host Innate Immunity. J Virol, 89, 10696-701.

TWU, K. Y., KUO, R. L., MARKLUND, J. & KRUG, R. M. 2007. The H5N1 influenza virus NS genes selected after 1998 enhance virus replication in mammalian cells. J Virol, 81, 8112-21.

TWU, K. Y., NOAH, D. L., RAO, P., KUO, R. L. & KRUG, R. M. 2006. The CPSF30 binding site on the NS1A protein of influenza A virus is a potential antiviral target. J Virol, 80, 3957-65.

URLAUB, H., RAKER, V. A., KOSTKA, S. & LUHRMANN, R. 2001. Sm protein-Sm site RNA interactions within the inner ring of the spliceosomal snRNP core structure. EMBO J, 20, 187-96.
VAN CAMPEN, H., EASTERDAY, B. C. & HINSHAW, V. S. 1989. Destruction of lymphocytes by a virulent avian influenza A virus. J Gen Virol, 70 (Pt 2), 467-72.

VAN DE SANDT, C. E., KREIJTZ, J. H. & RIMMELZWAAN, G. F. 2012. Evasion of influenza A viruses from innate and adaptive immune responses. *Viruses,* 4**,** 1438-76.

VAN REETH, K. 2000. Cytokines in the pathogenesis of influenza. *Vet Microbiol,* 74**,** 109-16.
VARGA, Z. T., GRANT, A., MANICASSAMY, B. & PALESE, P. 2012. Influenza virus protein PB1-F2 inhibits the induction of type I interferon by binding to MAVS and decreasing mitochondrial membrane potential. *J Virol,* 86**,** 8359-66.

VARGA, Z. T., RAMOS, I., HAI, R., SCHMOLKE, M., GARCIA-SASTRE, A., FERNANDEZ-SESMA, A. & PALESE, P. 2011. The influenza virus protein PB1-F2 inhibits the induction of type I interferon at the level of the MAVS adaptor protein. *PLoS Pathog,* 7**,** e1002067.

VERGARA-ALERT, J., BUSQUETS, N., BALLESTER, M., CHAVES, A. J., RIVAS, R., DOLZ, R., WANG, Z., PLESCHKA, S., MAJO, N., RODRIGUEZ, F. & DARJI, A. 2014. The NS segment of H5N1 avian influenza viruses (AIV) enhances the virulence of an H7N1 AIV in chickens. Vet Res, 45, 7.

VERSTREPEN, L., ADIB-CONQUY, M., KREIKE, M., CARPENTIER, I., ADRIE, C., CAVAILLON, J. M. & BEYAERT, R. 2008. Expression of the NF-kappaB inhibitor ABIN-3 in response to TNF and toll-like receptor 4 stimulation is itself regulated by NF-kappaB. J Cell Mol Med, 12, 316-29.
VIJAYAKRISHNAN, S., LONEY, C., JACKSON, D., SUPHAMUNGMEE, W., RIXON, F. J. & BHELLA, D. 2013. Cryotomography of budding influenza A virus reveals filaments with diverse morphologies that mostly do not bear a genome at their distal end. *PLoS Pathog,* 9**,** e1003413.

VIRMANI, N., BERA, B. C., SINGH, B. K., SHANMUGASUNDARAM, K., GULATI, B. R., BARUA, S., VAID, R. K., GUPTA, A. K. & SINGH, R. K. 2010. Equine influenza outbreak in India (2008-09): virus isolation, sero-epidemiology and phylogenetic analysis of HA gene. *Vet Microbiol,* 143**,** 224-37.

VON EYSS, B., MAASKOLA, J., MEMCZAK, S., MOLLMANN, K., SCHUETZ, A., LODDENKEMPER, C., TANH, M. D., OTTO, A., MUEGGE, K., HEINEMANN, U., RAJEWSKY, N. & ZIEBOLD, U. 2012. The SNF2-like helicase HELLS mediates E2F3-dependent transcription and cellular transformation. EMBO J, 31, 972-85.

VON MAGNUS, P. 1951. Propagation of the PR8 strain of influenza A virus in chick embryos. III. Properties of the incomplete virus produced in serial passages of undiluted virus. *Acta Pathol Microbiol Scand,* 29**,** 157-81.

WADA, N., MATSUMURA, M., OHBA, Y., KOBAYASHI, N., TAKIZAWA, T. & NAKANISHI, Y. 1995. Transcription stimulation of the Fas-encoding gene by nuclear factor for interleukin-6 expression upon influenza virus infection. J Biol Chem, 270, 18007-12.

WADDELL, G. H., TEIGLAND, M. B. & SIGEL, M. M. 1963. A New Influenza Virus Associated with Equine Respiratory Disease. *J Am Vet Med Assoc,* 143**,** 587-90.

WAFFARN, E. E. & BAUMGARTH, N. 2011. Protective B cell responses to flu--no fluke! *J Immunol,* 186**,** 3823-9.

WAHLE, E. & KELLER, W. 1996. The biochemistry of polyadenylation. *Trends Biochem Sci,* 21**,** 247-50.

WAKAHARA, T., KUSU, N., YAMAUCHI, H., KIMURA, I., KONISHI, M., MIYAKE, A. & ITOH, N. 2007. Fibin, a novel secreted lateral plate mesoderm signal, is essential for pectoral fin bud initiation in zebrafish. Dev Biol, 303, 527-35.

WALKER, J. K. L., THERIOT, B. S., GHIO, M., TREMPUS, C. S., WONG, J. E., MCQUADE, V. L., LIANG, J., JIANG, D., NOBLE, P. W., GARANTZIOTIS, S., KRAFT, M. & INGRAM, J. L. 2017. Targeted HAS2 Expression Lessens Airway Responsiveness in Chronic Murine Allergic Airway Disease. Am J Respir Cell Mol Biol.

WANG, B. X., WEI, L., KOTRA, L. P., BROWN, E. G. & FISH, E. N. 2017. A Conserved Residue, Tyrosine (Y) 84, in H5N1 Influenza A Virus NS1 Regulates IFN Signaling Responses to Enhance Viral Infection. Viruses, 9.

WANG, J., CAMPBELL, I. L. & ZHANG, H. 2008. Systemic interferon-alpha regulates interferon-stimulated genes in the central nervous system. Mol Psychiatry, 13, 293-301.

WANG, W., MU, X., ZHAO, L., WANG, J., CHU, Y., FENG, X., FENG, B., WANG, X., ZHANG, J. & QIAO, J. 2015. Transcriptional response of human umbilical vein endothelial cell to H9N2 influenza virus infection. Virology, 482, 117-27.

WANG, W., RIEDEL, K., LYNCH, P., CHIEN, C. Y., MONTELIONE, G. T. & KRUG, R. M. 1999. RNA binding by the novel helical domain of the influenza virus NS1 protein requires its dimer structure and a small number of specific basic amino acids. RNA, 5, 195-205.

WANG, X., BASLER, C. F., WILLIAMS, B. R., SILVERMAN, R. H., PALESE, P. & GARCIA-SASTRE, A. 2002. Functional replacement of the carboxy-terminal two-thirds of the influenza A virus NS1 protein with short heterologous dimerization domains. *Journal of virology,* 76, 12951-62.

WANG, X., HINSON, E. R. & CRESSWELL, P. 2007. The interferon-inducible protein viperin inhibits influenza virus release by perturbing lipid rafts. *Cell Host Microbe,* 2, 96-105.

WANG, X., LI, M., ZHENG, H., MUSTER, T., PALESE, P., BEG, A. A. & GARCIA-SASTRE, A. 2000. Influenza A virus NS1 protein prevents activation of NF-kappaB and induction of alpha/beta interferon. *J Virol,* 74, 11566-73.

WASIK, B. R., BARNARD, K. N., OSSIBOFF, R. J., KHEDRI, Z., FENG, K. H., YU, H., CHEN, X., PEREZ, D. R., VARKI, A. & PARRISH, C. R. 2017. Distribution of O-Acetylated Sialic Acids among Target Host Tissues for Influenza Virus. *mSphere,* 2.

WATTRANG, E., JESSETT, D. M., YATES, P., FUXLER, L. & HANNANT, D. 2003. Experimental infection of ponies with equine influenza A2 (H3N8) virus strains of different pathogenicity elicits varying interferon and interleukin-6 responses. Viral Immunol, 16, 57-67.

WEBBY, R. J. & WEBSTER, R. G. 2001. Emergence of influenza A viruses. Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 356, 1817-28.
WEBSTER, R. G., BEAN, W. J., GORMAN, O. T., CHAMBERS, T. M. & KAWAOKA, Y. 1992. Evolution and ecology of influenza A viruses. *Microbiol Rev,* 56, 152-79.

WHITNEY, I. E., KAUTZMAN, A. G. & REESE, B. E. 2015. Alternative splicing of the LIM-homeodomain transcription factor Isl1 in the mouse retina. Mol Cell Neurosci, 65, 102-13.
WIETECHA, M. S., CHEN, L., RANZER, M. J., ANDERSON, K., YING, C., PATEL, T. B. & DIPIETRO, L. A. 2011. Sprouty2 downregulates angiogenesis during mouse skin wound healing. Am J Physiol Heart Circ Physiol, 300, H459-67.

WISE, H. M., FOEGLEIN, A., SUN, J., DALTON, R. M., PATEL, S., HOWARD, W., ANDERSON, E. C., BARCLAY, W. S. & DIGARD, P. 2009. A complicated message: Identification of a novel PB1-related protein translated from influenza A virus segment 2 mRNA. *J Virol,* 83, 8021-31.

WOODWARD, A. L., RASH, A. S., BLINMAN, D., BOWMAN, S., CHAMBERS, T. M., DALY, J. M., DAMIANI, A., JOSEPH, S., LEWIS, N., MCCAULEY, J. W., MEDCALF, L., MUMFORD, J., NEWTON, J. R., TIWARI, A., BRYANT, N. A. & ELTON, D. M. 2014. Development of a surveillance scheme for equine influenza in the UK and characterisation of viruses isolated in Europe, Dubai and the USA from 2010-2012. *Vet Microbiol,* 169, 113-27.

XI, Y., KIM, T., BRUMWELL, A. N., DRIVER, I. H., WEI, Y., TAN, V., JACKSON, J. R., XU, J., LEE, D. K., GOTTS, J. E., MATTHAY, M. A., SHANNON, J. M., CHAPMAN, H. A. & VAUGHAN, A. E. 2017. Local lung hypoxia determines epithelial fate decisions during alveolar regeneration. Nat Cell Biol, 19, 904-914.

XU, H., LUO, J., MA, G., ZHANG, X., YAO, D., LI, M. & LOOR, J. J. 2017. Acyl-CoA synthetase short-chain family member 2 (ACSS2) is regulated by SREBP-1 and plays a role in fatty acid synthesis in caprine mammary epithelial cells. J Cell Physiol.

XU, K., KLENK, C., LIU, B., KEINER, B., CHENG, J., ZHENG, B. J., LI, L., HAN, Q., WANG, C., LI, T., CHEN, Z., SHU, Y., LIU, J., KLENK, H. D. & SUN, B. 2011. Modification of nonstructural protein 1 of influenza A virus by SUMO1. *J Virol,* 85, 1086-98.

XU, L. G., WANG, Y. Y., HAN, K. J., LI, L. Y., ZHAI, Z. & SHU, H. B. 2005. VISA is an adapter protein required for virus-triggered IFN-beta signaling. *Mol Cell,* 19, 727-40.

XU, Y., CHEN, X. & LI, Y. 2005. Ercc6l, a gene of SNF2 family, may play a role in the teratogenic action of alcohol. Toxicol Lett, 157, 233-9.

YAMANAKA, T., NIWA, H., TSUJIMURA, K., KONDO, T. & MATSUMURA, T. 2008. Epidemic of equine influenza among vaccinated racehorses in Japan in 2007. *J Vet Med Sci,* 70, 623-5.
YAMAZAKI, T. & ICHINOHE, T. 2014. Inflammasomes in antiviral immunity: clues for influenza vaccine development. *Clin Exp Vaccine Res,* 3, 5-11.
YANO, M., OUCHIDA, M., SHIGEMATSU, H., TANAKA, N., ICHIMURA, K., KOBAYASHI, K., INAKI, Y., TOYOOKA, S., TSUKUDA, K., SHIMIZU, N. & SHIMIZU, K. 2004. Tumor-specific exon creation of the HELLS/SMARCA6 gene in non-small cell lung cancer. Int J Cancer, 112, 8-13.

YE, L., GUO, L., HE, Z., WANG, X., LIN, C., ZHANG, X., WU, S., BAO, Y., YANG, Q., SONG, L. & LIN, H. 2016. Upregulation of E2F8 promotes cell proliferation and tumorigenicity in breast cancer by modulating G1/S phase transition. Oncotarget, 7, 23757-71.

YEE, M., DOMM, W., GELEIN, R., BENTLEY, K. L., KOTTMANN, R. M., SIME, P. J., LAWRENCE, B. P. & O'REILLY, M. A. 2017. Alternative Progenitor Lineages Regenerate the Adult Lung Depleted of Alveolar Epithelial Type 2 Cells. Am J Respir Cell Mol Biol, 56, 453-464.

YONDON, M., HEIL, G. L., BURKS, J. P., ZAYAT, B., WALTZEK, T. B., JAMIYAN, B. O., MCKENZIE, P. P., KRUEGER, W. S., FRIARY, J. A. & GRAY, G. C. 2013. Isolation and characterization of H3N8 equine influenza A virus associated with the 2011 epizootic in Mongolia. Influenza Other Respir Viruses, 7, 659-65.

YONEYAMA, M., KIKUCHI, M., NATSUKAWA, T., SHINOBU, N., IMAIZUMI, T., MIYAGISHI, M., TAIRA, K., AKIRA, S. & FUJITA, T. 2004. The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses. *Nat Immunol,* 5, 730-7.

YOO, J. E., PARK, Y. N. & OH, B. K. 2014. PinX1, a telomere repeat-binding factor 1 (TRF1)-interacting protein, maintains telomere integrity by modulating TRF1 homeostasis, the process in which human telomerase reverse Transcriptase (hTERT) plays dual roles. J Biol Chem, 289, 6886-98.

YOON, D. K., JEONG, C. H., JUN, H. O., CHUN, K. H., CHA, J. H., SEO, J. H., LEE, H. Y., CHOI, Y. K., AHN, B. J., LEE, S. K. & KIM, K. W. 2007. AKAP12 induces apoptotic cell death in human fibrosarcoma cells by regulating CDKI-cyclin D1 and caspase-3 activity. Cancer Lett, 254, 111-8.
YORK, A. & FODOR, E. 2013. Biogenesis, assembly, and export of viral messenger ribonucleoproteins in the influenza A virus infected cell. *RNA Biol,* 10, 1274-82.

ZENG, W., SUN, L., JIANG, X., CHEN, X., HOU, F., ADHIKARI, A., XU, M. & CHEN, Z. J. 2010. Reconstitution of the RIG-I pathway reveals a signaling role of unanchored polyubiquitin chains in innate immunity. *Cell,* 141, 315-30.

ZHANG, Y., MAO, D., ROSWIT, W. T., JIN, X., PATEL, A. C., PATEL, D. A., AGAPOV, E., WANG, Z., TIDWELL, R. M., ATKINSON, J. J., HUANG, G., MCCARTHY, R., YU, J., YUN, N. E., PAESSLER, S., LAWSON, T. G., OMATTAGE, N. S., BRETT, T. J. & HOLTZMAN, M. J. 2015. PARP9-DTX3L ubiquitin ligase targets host histone H2BJ and viral 3C protease to enhance interferon signaling and control viral infection. Nat Immunol, 16, 1215-27.

ZHAO, C., COLLINS, M. N., HSIANG, T. Y. & KRUG, R. M. 2013. Interferon-induced ISG15 pathway: an ongoing virus-host battle. *Trends Microbiol,* 21, 181-6.

ZHAO, Y., SILBAJORIS, R. & YOUNG, S. L. 1996. Identification and developmental expression of two activin receptors in baboon lung. Biochem Biophys Res Commun, 229, 50-7.

ZHOU, Z., HAMMING, O. J., ANK, N., PALUDAN, S. R., NIELSEN, A. L. & HARTMANN, R. 2007. Type III interferon (IFN) induces a type I IFN-like response in a restricted subset of cells through signaling

pathways involving both the Jak-STAT pathway and the mitogen-activated protein kinases. *J Virol*, 81, 7749-58.

ZIELECKI, F., SEMMLER, I., KALTHOFF, D., VOSS, D., MAUEL, S., GRUBER, A. D., BEER, M. & WOLFF, T. 2010. Virulence determinants of avian H5N1 influenza A virus in mammalian and avian hosts: role of the C-terminal ESEV motif in the viral NS1 protein. J Virol, 84, 10708-18.